

Winter 2017

A study of mathematics achievement, placement, and graduation of engineering students

Sara Hahler Blazek

Follow this and additional works at: <https://digitalcommons.latech.edu/dissertations>



Part of the [Applied Statistics Commons](#), [Higher Education Commons](#), and the [Other Engineering Commons](#)

**A STUDY OF MATHEMATICS ACHIEVEMENT,
PLACEMENT, AND GRADUATION OF
ENGINEERING STUDENTS**

by

Sara Hahler Blazek, B.S, M.S.

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

COLLEGE OF ENGINEERING AND SCIENCE
LOUISIANA TECH UNIVERSITY

February 2017

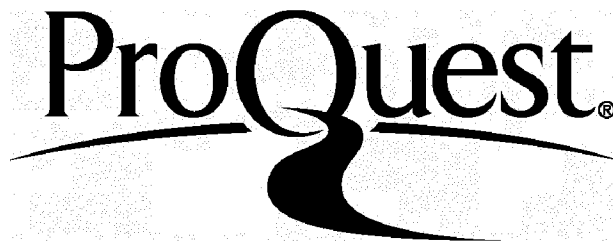
ProQuest Number: 10612784

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10612784

Published by ProQuest LLC(2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

LOUISIANA TECH UNIVERSITY

THE GRADUATE SCHOOL

November 8th, 2016

Date

We hereby recommend that the dissertation prepared under our supervision by
Sara Hahler Blazek entitled A Study of Mathematics
Achievement, Placement, and Graduation of
Engineering Students be accepted in
partial fulfillment of the requirements for the Degree of
Doctor of Philosophy

Maria K. Orr
Weizhong Dai Supervisor of Dissertation Research
Head of Department
CAM
Department

Recommendation concurred in:

U. Madadi
K. A. Cwan
D. Baring

Advisory Committee

Approved:

[Signature]
Director of Graduate Studies

Approved:

Sheryl S. Shoemaker
Dean of the Graduate School

Moham Hegab / SD
Dean of the College

ABSTRACT

The purpose of this study was to determine how background knowledge impacts freshmen engineering students' success at Louisiana Tech University in terms of grades in two different freshman classes and graduation. To determine what factors impact students, three different studies were implemented. The first study used linear regression to analyze which demographic and academic variables significantly impacted freshman math and engineering courses. Using regression discontinuity, the second study determined if the university's placement requirement for Pre-Calculus was appropriate. The final study analyzed factors that impact graduation for engineering students as well as other disciplines to determine which significant variables were unique to engineering.

Numerous studies have focused on factors that influence engineering students' first year retention and graduation. However, studies have reached various conclusions which appear contradictory. For example, some studies determined that sex was a significant influence while other studies determined that it was not influential. Multiple studies found that academic factors such as high school rank or grade point average were important for engineering student success. With conflicting results, it is important to determine what is true for engineering students at Louisiana Tech University. Identifying factors that influence first-year grades, mathematics placement, and graduation could be useful for the recruitment, retention, and academic success of students.

The participants in this study were freshmen engineering students who were enrolled in Louisiana Tech University between the fall of 2006 and the fall of the 2014/2015 school year. The variables used in the studies included high school GPA, ACT component scores, race, state residency, sex, enrollment in either the integrated freshmen engineering sequence or Living with the Lab, and peer economic status.

APPROVAL FOR SCHOLARLY DISSEMINATION

The author grants to the Prescott Memorial Library of Louisiana Tech University the right to reproduce, by appropriate methods, upon request, any or all portions of this Thesis. It is understood that “proper request” consists of the agreement, on the part of the requesting party, that said reproduction is for his personal use and that subsequent reproduction will not occur without written approval of the author of this Thesis. Further, any portions of the Thesis used in books, papers, and other works must be appropriately referenced to this Thesis.

Finally, the author of this Thesis reserves the right to publish freely, in the literature, at any time, any or all portions of this Thesis.

Author Sara Hahle Blyde

Date Jan. 23rd, 2017

DEDICATION

To my mother, my father, and my grandmother – from the beginning to the end,
their love, encouragement, and godly guidance have enabled me to finish the race.

TABLE OF CONTENTS

ABSTRACT	iii
DEDICATION	vi
LIST OF TABLES	xi
LIST OF FIGURES.....	xiii
ACKNOWLEDGEMENTS	xiv
CHAPTER 1 INTRODUCTION	1
1.1 Research Need and Importance	1
1.2 Purpose of Study.....	2
1.3 Theoretical Framework.....	5
1.4 Dissertation Overview	6
CHAPTER 2 BACKGROUND	8
2.1 Overview	8
2.2 Aspects that Affect Graduation	8
2.3 Importance of the First Year.....	10
2.4 College Mathematics	12
2.5 Placement	14

2.6	Institution Specifics	16
CHAPTER 3 GENERAL INFORMATION		19
3.1	Research Questions	19
3.2	Data Sources and Analyses.....	20
3.3	Predictor Variables	22
3.3.1	Explanation of Variables	22
3.3.2	Transformation of Variables	25
3.3.3	Justification	26
3.4	Outcome Variables	28
3.4.1	Achievement Study	28
3.4.2	Placement Study	28
3.4.3	Graduation Study.....	29
3.5	Analysis Techniques.....	29
3.5.1	Linear Regression.....	29
3.5.2	Regression Discontinuity	33
3.5.3	Logistic Regression	36
3.5.4	Confidence Intervals and Significance.....	39
3.6	Protection of Human Subjects	40
CHAPTER 4 STUDY ONE: ACHIEVEMENT		43

4.1	Approach and Justification	43
4.2	Predictor and Outcome Variables.....	44
4.3	Participants	44
4.4	Method.....	45
4.5	Expected Outcomes	46
4.6	Limitations.....	46
4.7	Results for Study One: Achievement	48
4.7.1	Pre-Calculus Models	51
4.7.2	Discussion of Pre-Calculus Results.....	57
4.7.3	Engineering Problem Solving I Models	60
4.7.4	Discussion of Engineering Course Results	62
CHAPTER 5 STUDY TWO: PLACEMENT		64
5.1	Approach and Justification	64
5.2	Predictor and Outcome Variables.....	66
5.3	Participants	67
5.4	Method.....	68
5.5	Expected Outcomes	70
5.6	Limitations.....	71
5.7	Results for Study Two: Placement	72

CHAPTER 6 STUDY THREE: GRADUATION.....	77
6.1 Approach and Justification	77
6.2 Predictor and Outcome Variables.....	78
6.3 Participants	80
6.4 Method.....	83
6.5 Expected Outcomes	83
6.6 Limitations.....	84
6.7 Results for Study Three: Graduation	85
CHAPTER 7 SUMMARY AND CONCLUSION	105
7.1 Study One: Achievement.....	105
7.2 Study Two: Placement.....	107
7.3 Study Three: Graduation	107
7.4 Future Work and Conclusion.....	109
APPENDIX A STUDY THREE: LIST OF MAJORS	110
APPENDIX B APPROVALS	113
B.1 IRB Approval Letter.....	114
B.2 IRB Proposal	117
BIBLIOGRAPHY	125

LIST OF TABLES

Table 1	Study One Predictor Variables and Their Ranges.....	26
Table 2	Definitions of Variables for First Stage Equation.....	34
Table 3	Ill and Healthy Patients Given New Drug or Placebo.....	37
Table 4	Definitions of Variables for HL Statistic Formula.....	38
Table 5	Definitions of Variables for Coefficient Confidence Interval.....	39
Table 6	Definitions of Variables for Coefficient Confidence Interval: Logistic Regression.....	40
Table 7	Study One Sample by Race, Sex, and State.....	45
Table 8	Correlation Coefficients above 0.6.....	49
Table 9	Linear Regressions with High School GPA and Rank Against Grade in Pre-Calculus.....	50
Table 10	ACT Math as a Predictor of Pre-Calculus and Engineering Grade.....	51
Table 11	Pre-Calculus Models.....	53
Table 12	Pre-Calculus Model with Interaction Term.....	54
Table 13	Pre-Calculus Model with Backward Elimination: All Variables Significant.....	55
Table 14	Pre-Calculus Models with Backward Elimination: Models 9-12.....	56
Table 15	Engineering Models for Achievement Study.....	61
Table 16	Engineering Model with Interaction Term.....	62

Table 17	Descriptive Statistics for Study Two Participants.....	68
Table 18	Descriptive Statistics for Placement Study Participants.....	70
Table 19	Academic Statistics for Placement Study Participants.....	70
Table 20	First Stage Regression Discontinuity: Final Three Models.....	74
Table 21	Second Stage Regression Discontinuity: Final Three Models.....	75
Table 22	Mean GPA, Rank, and ACT by Major.....	82
Table 23	Study Three Sample by Race and Sex.....	83
Table 24	Correlation Coefficients between High School GPA and Rank.....	86
Table 25	Condition Number for Each Major Type.....	86
Table 26	High School GPA and Rank Models.....	87
Table 27	Engineering Models.....	89
Table 28	Engineering Model for ACT Reading.....	91
Table 29	Average ACT Scores by Race for Engineering Students.....	91
Table 30	Psychology Models.....	92
Table 31	Social Sciences Models.....	94
Table 32	Health Models.....	95
Table 33	Business Models.....	96
Table 34	Business Models (Stepwise).....	96
Table 35	Education Models.....	97
Table 36	Final Graduation Models.....	101
Table 37	Linear Regression for Graduation versus LWTL Variable.....	103
Table 38	Engineering Models Including the LWTL Variable.....	104
Table 39	Majors within Each Group in Study Three.....	111

LIST OF FIGURES

Figure 1	Regression lines of groups by grade in Pre-Calculus and ACT math score (cutoff score = 26). For the figure, $f(x)$ = students with ACT math scores of 24 and 25, $h(x)$ = students with ACT math scores between 26 and 28, and $g(x)$ = projected grades for the second group. Also, recall that grade in Pre-Calculus was transformed into numbers; therefore, 4 represents an “A,” 3 a “B,” and so on	73
----------	---	----

ACKNOWLEDGEMENTS

This dissertation represents years of hard work that would not have been possible without the help and support from colleagues, family, and friends. I would like to thank Dr. Galen Turner for encouraging me to apply for this program and for all the mentorship provided during the journey. Thanks to Tara, Galen, and Ben for providing a home away from home.

I also appreciate and recognize the time and effort my advisor, Dr. Marisa Orr, spent helping me as well as all my committee members. To Carrie Falke Maggio – thank you for all the notes you provided for me, for being a study partner when needed, and the friendship. I would also like to thank this special group of friends for their support, notes, and guidance throughout the years – Dr. Ankunda Kiremire, Dr. David Irakiza, and Alicia Kiremire. I wouldn't have made it through grad school without you guys.

To Dr. Krystal Corbett – thank you for your advice, your notes, your listening ear, opening up your home, and most importantly for your friendship. I am also grateful to Tiffany, Torri, Taylor, Trinity, Justina, and Dani for their continuous support throughout my life. Josh, thank you for being an awesome brother who always has my back and always wants the best for me. I appreciate the Blazek family, especially Mr. Tom and Mrs. Cindy, for their overwhelming encouragement to keep the faith. James and Brooke – thanks for all the times you cooked for us and for providing us with such

excellent company so often. I would like to thank Robbie Ritchie for always pushing me to do my best, but being proud of me no matter what.

Dad, thank you for encouraging me to never settle and for giving the absolute best advice no matter the circumstance.

Mom, thank you for proofreading all my work! I appreciate your never-failing belief that I can accomplish whatever I set my mind to.

Daniel, thank you for your everyday support – through the lows and highs, I can always count on you to encourage me and also make me smile.

Thank you God for placing these people in my life and allowing me this opportunity.

CHAPTER 1

INTRODUCTION

1.1 Research Need and Importance

Throughout the nation, the need for engineers is growing. In fact, the number of STEM (Science, Technology, Engineering, and Mathematics) jobs – of which engineering jobs accounted for one-third as of 2010 – within the United States has tripled during the last few years and the growth is projected to continue until at least 2018 [1]. However, there are not enough engineers to meet the demands of the developing workforce [2]. As of 2012, bachelor degrees in engineering disciplines make up less than five percent of all bachelor degrees awarded in the nation [3]. Furthermore, the number of engineering students graduating is not increasing.

A study published in 2010 indicated that the number of STEM graduates, as a whole, has declined in recent years [4]. Other studies have found that engineering retention to graduation has remained relatively stagnant [5, 6, 7]. Additionally, the number of incoming freshmen majoring in engineering has been stagnant in recent years [8]. Researchers have also found that the migration rate of students switching into engineering is very low [9].

While the high demand and low supply of engineers is a national concern, many universities have decided to help combat the issue. Most universities approach the problem by focusing on two particular aspects, student recruitment and freshman

retention, with the ultimate goal of increasing engineering graduation rates. If more students are recruited, then the pool of students that are retained after the first year may grow. Similarly, if there are more retained students through the first year then there is a higher likelihood of students graduating with an engineering degree. After reaching this conclusion, the first question researchers must answer is what type of students should be recruited. If the final goal is to increase the number of students graduating in engineering, then it follows that students more likely to graduate should be recruited. In this vein, multiple institutions have started to conduct research on what factors, such as high school data or demographic variables, have a strong relationship (whether it is negative or positive) to retention through the first year and persistence to graduation. Although much research has been undertaken in this area, different studies have indicated different results. Some studies state that high school academics such as GPA or rank significantly influence an engineering major's first year or graduation while other researchers make a case for standardized test scores, race, or sex being important variables to consider when predicting engineering student success [10, 11, 12, 13].

What is true for undergraduate engineering students at Louisiana Tech University? With a growing and nationally known engineering program, it is important that data from this specific university be analyzed so that the results can best assist the program in recruiting, retaining, and graduating engineering students by making conclusions directly derived from the university's student data.

1.2 Purpose of Study

It is understood that there is a need for more engineers and therefore for increased recruitment and retention in engineering programs across the nation in hopes of

producing more engineering graduates. One way to address these pressing concerns is to study current engineering students' ability to succeed, academically speaking, in an engineering program; more specifically, to study factors that negatively or positively influence the freshmen year (a crucial point in retaining students to graduation) along with retention to graduation. By studying these topics, the information gathered can then be used for both recruiting (to identify students who show great potential) and freshman retention (through better placement).

All of this information contributes to the primary goal of this research: to determine information to assist Louisiana Tech University in graduating more engineers and consequently help engineering students graduate. To do this, eight particular research questions were answered through three individual studies. Using linear regression to analyze freshmen math and engineering grades, the first study (the achievement study) focused on the following:

1. Is *ACT math score* a significant influence on the final grades of Pre-Calculus for engineering students at Louisiana Tech University?
2. Do other variables, besides *ACT math score*, have a significant relationship with the final grades of Pre-Calculus for engineering students at Louisiana Tech University?
3. Is *ACT math score* a significant influence on the final grades of Engineering Problem Solving I for engineering students at Louisiana Tech University?
4. Do the other variables, besides *ACT math score*, have a significant relationship with the final grades for Engineering Problem Solving I?

The second study (the placement study) used a regression discontinuity model to analyze the differences between engineering students required to take remedial math courses and those not required to take remedial courses in terms of grade in Pre-Calculus.

5. Is the ACT cutoff score used to place students in Pre-Calculus at Louisiana Tech University one such that students right above and below the cutoff do similarly well in Pre-Calculus?

The last study (the graduation study) implemented logistic regression to analyze factors that influence graduation for engineering majors as well as other disciplines in order to answer these questions:

6. What factors are influential to graduation for engineering students at Louisiana Tech University?
7. What factors that influence graduation are unique to engineering students at Louisiana Tech University?
8. Does enrollment in *Living with the Lab* have an influence on graduation for engineering students?

A detailed explanation of *Living with the Lab* is given on page 22 and 23.

The combination of these studies provided cumulative information that can be used to help engineering students succeed academically. The first study indicated what background knowledge or demographic factors are influential to freshmen engineering students' grade in the first math and engineering classes required for Louisiana Tech's engineering program; this information can be used to identify high school students who may be likely to succeed in engineering. The second study determined if students are being correctly placed in this first mathematics class. These two pieces are important as

the first year in an engineering program, specifically freshmen GPA, has a large amount of influence not only on freshman retention but also on graduation [14, 15, 16]. The third study determined what high school, demographic, and various college academic factors have an impact on graduation. The results from this study can help Louisiana Tech University student success specialists identify students in the engineering program who may be at risk of not graduating with an engineering degree.

1.3 Theoretical Framework

Constructivism is a theory of knowledge largely attributed to scholars like Piaget and Vygotsky. The theory argues that humans construct knowledge from an interaction between experiences and ideas already formed [17, 18, 19]. More specific to education and learning, constructivism offers a different view of learning by defining it as a process of constructing meaning from a personal viewpoint [20]. A major principle of this learning theory is that new knowledge is created on the basis of background or prior knowledge.

A basic definition of background knowledge is “the raw material that conditions learning” [21]. Others define it as the knowledge and skills a person already knows or has about a certain concept or “all knowledge learners have when entering a learning environment that is potentially relevant for acquiring new knowledge” [22, 23, 24]. Thus, a fundamental assumption throughout these studies is that the acquisition of new knowledge is facilitated by the knowledge that a person already has. This is evident for students beginning courses at a university – a professor assumes that students have a certain level of knowledge when beginning classes. In fact, many studies have determined that prior knowledge has a significant influence on a student’s academic

success [25, 26, 27]. The more knowledge a student has about a topic, the better prepared they are to build upon the topic. Utilizing this idea, the studies in this dissertation will use measures of background knowledge, as well as variables that have been shown in other studies to be significant for engineering academic success, with the aim to identify which variables are most influential for the success of engineering students at Louisiana Tech University.

1.4 Dissertation Overview

The layout of this dissertation is as follows. First, a review of relevant literature will be presented including aspects that affect graduation for engineering students and the first year at a university as well as institution specifics (Chapter 2). This will include topics such as aspects that affect graduation, the importance of the first year academically and mathematical preparation, and placement policies followed by institution specifics (the institution from which the data for this work was retrieved). In Chapter 3, an explanation of general information is given, such as the research questions of the dissertation and data sources which are common to the three studies that are a part of this work. Chapter 3 also generally explains the predictor and outcome variables used in each study while specific variable details are outlined in later chapters. The three analysis techniques used in this work are linear regression, fuzzy regression discontinuity, and logistic regression for the three separate studies that make up the dissertation work (each explained in detail in Chapter 3 Section 5). Next, the specific details of study one (the achievement study) are relayed: approach, justification, variables, method, expected outcomes, limitations, and results. Each of these areas are explained in detail in Chapter 4 section by section. Details of study two (the placement

study) and then study three (the graduation study) follow in Chapter 5 and 6 respectively.

A summary and conclusion of each study is provided followed by a discussion of possible future work in Chapter 7.

CHAPTER 2

BACKGROUND

2.1 Overview

Considering the three different studies presented in this dissertation, multiple topics are addressed in the literature review. A primary goal of this research study is to determine information to assist Louisiana Tech University in graduating more engineers. Therefore, a section of the literature review will be devoted to retention to graduation for engineering students. The first year at a university, particularly in regards to math courses, is also important to the success of engineering students. Hence, aspects that influence an engineering student's first year and the importance of mathematics will also be reviewed. Another section will review the influence of remedial classes on student academic success and how students are placed in these classes. Lastly, institution specifics, which influences the data used in the study, are discussed.

2.2 Aspects that Affect Graduation

What factors influence engineering students in terms of graduating? A study sampling business, education, and STEM majors to determine factors that influence graduation highlighted the fact that "students with higher high school rankings, no matter their race..., gender... or which school district they are coming from, should be encouraged into STEM majors" [10]. Scott, Tolson, and Huang, studying math and science majors, discovered that a student's high school rank and combined mathematics

and reading SAT scores positively correlated with retention to the junior year [28]. Another study found that both *high school GPA* and *math SAT* were positively correlated with graduation rates for each of the six universities tested in the study, along with factors such as *sex*, *ethnicity*, and *citizenship* for some of the universities [13]. Using a different standardized test in their study, Moller-Wong and Eide stated that *ACT math score* and *high school rank* among other variables were statistically significant when determining if students would stay in engineering until graduation [12]. Besides these factors, the study indicated several more variables were positive and significant: high number of transfer hours, non-residents of the state in which the school was located, and high numbers of high school physics and social science classes. On the opposite end, Moller-Wong and Eide also found that being African American or having an ACT composite score of 35 or 36 puts a student at risk for not graduating with an engineering degree. French, Immekus, and Oakes used hierarchical logistical and linear regression to determine factors that influenced engineering student success and persistence [11]. Their results indicated that *SAT math score* and *high school rank* were positive predictors for a student's cumulative GPA (after 8 semesters) while *cumulative GPA* was the only significant variable in predicting persistence to the eighth semester.

Studying the variable *sex*, other researchers have indicated that females are slightly more likely to defect from engineering regardless of race [29, 30, 31]. Comparing engineering majors to other majors, another study concluded that females were less likely to persist to the senior year [32]. Other results indicated that students who persisted spent less time working off campus and more time preparing for class. Multiple studies

connected persistence and achievement in engineering with students' self-efficacy [33, 34].

Besides high school measures and self-efficacy, studies have also concluded that success during the first year of college, particularly freshmen GPA and achievement in mathematics courses, is a strong predictor of graduation for engineering students [14, 15]. Research has also shown that placement in the appropriate mathematics course during the freshman year influences academic success for engineering majors and therefore also indirectly influences graduation [35, 36, 37].

2.3 Importance of the First Year

In this section, research studies on the importance of the first year at a university will be reviewed. An engineering student's first year at a university plays a crucial role in determining the student's future academic career. It has been shown that more students leave engineering between freshmen and sophomore years than any other period of time in the college experience [16]. Suresh found that students who did well in first-year courses had a high tendency to complete an engineering program [38]. Lebold and Ward discovered that first and second semester grades were strong predictors of engineering persistence [14]. Another study evaluated Purdue University's freshmen engineering students. Using longitudinal data, the results indicated that student retention to the sixth semester was higher for students with a higher first semester GPA [15]. Yoon Yoon, Imbrie, and Reed studied first-year engineering students, mathematics courses, and graduation at Texas A&M University [39]. The findings from the study showed that students with AP or CLEP exam credits for Calculus I had a higher chance of graduating in engineering than students who took Calculus I at the university. Likewise, Moses *et al.*

used logistic regression to show that *calculus readiness*, as well as *high school GPA*, was predictive of first-year retention for engineering majors at East Carolina University [40].

Based on the literature, a successful academic year of a freshman student is extremely important to retaining students until graduation. Knowing that grades in the first year have a significant influence on graduation, the next step is to discover what factors have an impact on freshmen grades.

Examining engineering majors as well as other types of students, multiple studies have reported that measures of high school performance and standardized test scores are related to first-year grades [41, 42, 43]. Patterson and Mattern found that *SAT component scores*, both math and verbal, and *high school GPA* were both strongly correlated with first-year GPA for first-time, first-year students from 160 institutions who enrolled in the fall of 2010 [44]. An earlier study by Camara and Echternacht concluded the same results: student's SAT scores and high school grades were significant predictors of a student's freshmen GPA, although *high school GPA* was more predictive than *SAT score* [45]. In their study, these results tended to hold true for all subgroups of students. Moses *et al.* used logistic regression analysis to find that *high school GPA* and *calculus readiness* were critical to first-year retention [40]. A study concerning freshmen at the University of Michigan indicated that *ACT math score* was a positive predictor of a passing grade in first-year engineering courses [46]. Analyzing freshmen of all majors at a single university, DeBerard, Spielmans, and Julka discovered that *sex* (being female), *high school GPA*, and *SAT score* were positively associated with students' first-year cumulative GPA [47].

Testing academic variables as well as other constructs, a study involving freshmen students in Australia established that high school academic performance along with being agreeable, being an introvert, and using self-regulated learning strategies were indicators of first-year, first-semester academic success [48]. Analyzing non-academic factors, Honken and Ralston performed a case study and concluded that having a parent who did not graduate from college made a student less likely to be retained to the second year [49]. On the other hand, the same study found that studying with fellow students meant one was more likely to be retained. Other studies have found that *self-efficacy* plays a role in achieving academically in the first year. Chemers, Hu, and Garcia completed a longitudinal study of first-year students and indicated that *academic self-efficacy* (and *optimism*) were strongly related to first-year academic performance [50]. A different study examined first-year college students and psychosocial factors and also found that *academic self-efficacy*, along with *organization* and *attention to study*, was predictive of freshman GPA in the first semester [51].

Another variable that affects engineering success is mathematical preparation [9, 11, 12, 52]. Unfortunately, a lack of preparation in this subject often prevents a student from succeeding academically.

2.4 College Mathematics

In the first year of an engineering program, a student may or may not be required to take an introductory engineering course. At Purdue University and other universities, engineering majors do not have the opportunity to take engineering classes until the second year [53]. Some engineering programs, like the one at Wright State, include an introductory math class that focuses on engineering applications [54]. At Louisiana Tech

University engineering majors take mathematics courses alongside freshman engineering classes. Regardless of the program or university, the majority of engineering majors are required to take multiple mathematics classes in the first year. However, research has indicated that “the biggest factor in the failure of (engineering) students is inadequate competence in mathematics” [37]. Other researchers concur, such as Moyo who states that engineering students at the University of Venda often struggle in first-year math classes [55]. Astin and Oseguera agree that math skill is a primary influence on retention, especially for engineering students [35]. Another study concluded that success in mathematics courses relates to a higher likelihood of retention to graduation [36]. Robinson, studying high school factors and STEM majors, found that students were more likely to be successful in college if they had taken advanced math and science classes in high school [56]. This was especially true for engineering majors. Other research has specifically shown that university freshman students majoring in an engineering discipline struggle with interpreting graphs, finding the equation of a line, calculating the volume of a prism or cylinder, and solving probability problems [57]. In fact, researchers state that a lack of mathematical preparedness is the largest cause of drop-outs in the freshmen year and has resulted in a decrease of the popularity of undergraduate engineering degrees [58, 59, 60, 61].

Though mathematics may be an obstacle to an engineering degree, it is also essential to the study of engineering. Kirschenman and Brenner define engineering as “the application of mathematics and sciences to building of projects for the use of society” [62]. A student unable to grasp mathematical concepts may also struggle in applying said concepts to engineering problems. Other researchers, Zeidmane and Sergejeva, state that

“[m]athematics studies have an impact on the development of the necessary outcomes for engineers both directly (mathematics serves as a tool for solving and calculating various problems) and indirectly (mathematics develops skills to formulate, solve engineering problems etc.)” [63]. Math is at the core of engineering - the concepts taught in the mathematics classes are the building blocks upon which engineering concepts are constructed [64].

If mathematics is a stumbling block for students, yet necessary for the completion of an engineering degree, then how do educators bridge this gap? One step is ensuring that students are initially placed in math courses that are appropriate for the student’s skill level.

2.5 Placement

It is obvious that placing a freshmen student in the appropriate math class is important. Ohland, Yuhasz, and Sill argue that starting college mathematics in the correct course (one for which for they are prepared) is important to retention as well academic success in a class [65]. Medhanie *et al.* states that placing students into “an appropriate college mathematics course is key to the success of students in a course” [66]. Additionally, it has been found that if a student is not adequately prepared for a class then it is unlikely that he or she will pass the class; on the other hand, if a student is over prepared for a class then it is possible he or she will underperform [67]. A student more likely to succeed in a math class is more likely to have a high freshman GPA and therefore have a greater chance of graduating with an engineering degree.

Universities have a variety of ways to determine which math class is best for a student to take upon entering the university. Many universities use some type of

placement exam to determine the class a student will enter. In fact, NCES reported that in 1995 sixty-four percent of all institutions in the nation used a placement exam to determine the need for mathematics remediation and in 2000 sixty-one percent of the institutions used a placement exam [68]. One such school is the University of Wisconsin-Milwaukee. Students take a placement exam, before the summer term, and if they place into a class below Calculus I, then the students participate in a summer bridge program that utilizes a computer-based math review [69]. At the University of Arizona, students could choose one of two placement tests to take (one that covered intermediate algebra skills and another that covered college algebra and trigonometry) and were then placed in an appropriate math class based on the test results [70]. A self-assessment test is given to all incoming students at the University of Sydney to assist them in deciding whether or not to enroll in the highest level math class available to freshmen [71]. The University of California and the California State University systems collaborated to create a placement test in their Mathematics Diagnostic Testing Project (MDTP) for use in placing students.

Other institutions use multiple criteria to place students. At St. Olaf College in Minnesota, a placement test coupled with multiple categories of student data, including high school rank and GPA, is used to place students [72]. For students at Cottey College in Missouri, three different measures are considered when matching them with a class: placement test score, standardized test scores, and the number of math courses taken in high school [73].

A number of universities use standardized test scores to place students and sometimes combine this measure with a placement test. For example, at a public university in Ohio, ACT is the primary tool to place students; however, in two different

cases, students are also given a placement exam [74]. The first exception is if a student has already taken a remedial course and the second is if a student has an ACT math score of 20 or below. In these situations, the student is then allowed to take the COMPASS placement test during orientation. A university studied by Jacobson institutes a similar practice – initial placement is decided by ACT score, but students are given a chance to increase this score by also taking the COMPASS test at the university [75]. Middle Tennessee State University, also uses either ACT, SAT, or COMPASS scores to place students [76].

At Louisiana Tech University, a public university and the focus institution of this study, a student's ACT math score is primarily used to place students. However, if a student scores a 24 or 25 then he or she is given the opportunity to pass a credit exam for College Algebra (usually taken within the first three days of class during the Fall term) and join the group of students taking Pre-Calculus.

2.6 Institution Specifics

The site selected for this study is Louisiana Tech University, the home institution of the researcher. Established in 1894, Louisiana Tech is a public university located in Ruston, Louisiana. It currently enrolls over 12,000 undergraduate and graduate students from 48 different states and 68 foreign countries. The university is a selective admissions research institution that awards bachelor, master, and doctoral degrees [77].

At Louisiana Tech University, the engineering disciplines available are: biomedical, chemical, civil, cyber, electrical, industrial, mechanical, and nanosystems. A cyber engineering undergraduate degree first became an option in the 2012-2013 school year. Another option for engineering majors was available from fall of the 2003-2004

school year to the summer of 2014 - basic engineering. This major was a placeholder major for first-year students who were undecided; it had to be changed to a specific engineering discipline by the beginning of a student's second year.

For all freshmen engineering majors, the suggested mathematics courses for the first year are Pre-Calculus, Calculus I, and Calculus II. These math classes are the corequisites for the freshman engineering courses: Engineering Problem Solving I, Engineering Problem Solving II, and Engineering Problem Solving III, respectively. To enroll in Pre-Calculus, a student must have credit for College Algebra (automatically given to students with ACT math scores 26 or above) and high school trigonometry (if the student did not have the class during high school, they are required to take Trigonometry at the university level). However, the vast majority of engineering students at Louisiana Tech self-report that they have taken trigonometry in high school.

Students with ACT math scores 25 and below are required to take one or more remedial classes or pass a placement exam. For this work, "remedial" will refer to a math class of lower level than Pre-Calculus. For example, engineering students with ACT math scores between 24 and 25 are given the option of enrolling in College Algebra or taking a credit exam to earn credit for that class. However, students having a score between 22 and 23, the student is required to enroll in College Algebra. Other requirements are given for ACT math scores below 22.

If a student is placed into remedial math classes during the first year, the student is not allowed to enroll in the Calculus sequence and is also not able to enroll in the co-requisite freshman engineering sequence. Students able to enroll in Pre-Calculus (or a higher level math class) in the first quarter of their freshmen year are considered to be

“on track to graduation” because they can typically graduate in four years. The option of enrolling in a higher level math class (Calculus I or II) became available to students with ACT math scores of 26 or above in the summer of 2011. To enroll directly in these classes, the student must first pass a placement exam; on average, six students a year succeed.

It should also be noted that Louisiana Tech University is on the quarter calendar, but uses the semester credit hour system. Thus, students attend classes each academic year in three ten-week quarters during Fall, Winter, and Spring. Since a full-time student has a minimum load of 24 semester credit hours, typically distributed over two semesters, a full-time student at Louisiana Tech must enroll in at least eight hours a quarter.

CHAPTER 3

GENERAL INFORMATION

3.1 Research Questions

To achieve the goal of determining information that will help Louisiana Tech University by assisting engineering majors at the university to graduate with an engineering degree, eight research questions were addressed through three individual studies. A reminder of these studies and the corresponding research questions are provided. Using linear regression to analyze freshmen math and engineering grades, the first study focused on the following:

1. Is *ACT math score* a significant influence on the final grades of Pre-Calculus for engineering students at Louisiana Tech University?
2. Do other variables, besides *ACT math score*, have a significant relationship with the final grades of Pre-Calculus for engineering students at Louisiana Tech University?
3. Is *ACT math score* a significant influence on the final grades of Engineering Problem Solving I for engineering students at Louisiana Tech University?
4. Do the other variables, besides *ACT math score*, have a significant relationship with the final grades for Engineering Problem Solving I?

A student's ACT math score is of particular interest as it is used to place students in their first math class at Louisiana Tech. The second study used a regression discontinuity

model to analyze the differences between engineering students required to take remedial math courses and those not required to take remedial courses in terms of the final grade in Pre-Calculus in order to answer:

5. Is the cutoff score, an ACT math score of 26, used to place students in Pre-Calculus at Louisiana Tech University one such that students right above and below the cutoff do similarly well in Pre-Calculus?

The last study implemented logistic regression to analyze factors that influence graduation for engineering majors as well as other disciplines in order to answer these questions:

6. What factors are influential to graduation for engineering students at Louisiana Tech University?
7. What factors that influence graduation are unique to engineering students at Louisiana Tech University?
8. Does enrollment in *Living with the Lab* have an influence on graduation for engineering students?

The remainder of this chapter addresses the data sources and predictor variables that are common to all studies (sections two and three) before the outcome variables and methods used in each study are described in more detail in sections four and five.

3.2 Data Sources and Analyses

The participants for all three studies were first-time-in-college (FTIC) freshmen at Louisiana Tech University. All academic and demographic data were retrieved from institutional records at Louisiana Tech University. To identify FTIC freshmen, only students enrolled in a required university seminar class (UNIV 100) were included in the

study. Transfer students are not required to take this class and were therefore not included in the study.

Race, ethnicity, and international status were confounded into one variable for this study due to limitations in the way the data were collected. The students were given the following options from which to choose on their admission application: White, Black/African American, Hispanic, Pacific Islander, Asian American, American, non-resident alien, and decline to identify. Hispanic, Pacific Islander, and Asian American students were excluded from two of the studies because their numbers were insufficient for generalizability. Non-resident aliens and those who declined to identify a race were excluded because their race was unknown. Therefore, study one and study three included only students who reported their race as Black or White. Thus, the findings may not hold for other groups.

Cohorts before 2006 were eliminated from the studies due to a change in the mathematics placement policy. From 2006 to 2016, the requirement remained the same—an ACT math score of 26 or above made a student eligible to enroll in Pre-Calculus if the student also self-reported having taken trigonometry in high school. The most recent academic data available at the time of the study included information up to the first quarter of the 2016 school year.

The second source of data is the National Center for Education Statistics (NCES). Free lunch enrollment data from the Common Core of Data (CCD) was used to calculate the peer economic status variable described in the next section. At the time of the study, the most recent data from the NCES on free lunch participation was for the 2014 cohort, so the sample window was reduced to 2006-2014.

To clean (replace or delete corrupted/inaccurate information) the data retrieved from these sources, two programs were used: Microsoft Excel® and R Studio. To perform the statistical analyses for each study in this work, R Studio was used.

3.3 Predictor Variables

3.3.1 Explanation of Variables

The predictor variables used were: *high school GPA*, *high school rank*, *sex*, *race*, *ACT math score*, *ACT English score*, *ACT reading score*, *ACT science score*, *state residency*, *peer economic status*, and enrollment in either the integrated freshman engineering sequence or *Living with the Lab* curriculum (LWTL). A summary table of the variables and their ranges are given on page 26. Most of the variables listed are self-explanatory; a more detailed explanation of the curriculum variable and peer economic status will now be given.

The integrated freshman engineering sequence referred to in this work was a fully integrated mathematics, science, and engineering curriculum. It was initially piloted in 1997 and fully implemented in the 1999-2000 school year. The integrated curriculum had a focus on connections across disciplines and reliance on cooperative learning. More information concerning this curriculum can be seen in [78], [79], and [80].

In the 2008-09 school year, a new iteration of the curriculum for the freshmen engineering courses (Engineering Problem Solving I, II, and III) was implemented for all students. The new design of the course focused on the same principles, but also strongly focused on projects and encouraged student ownership using two main platforms throughout the years, the Parallax Boe-Bot and the Arduino; this new design was called *Living with the Lab* (LWTL). More information concerning the details and material for

this sequence can be found in [81], [82] and [83]. With such a significant change in the program, for the purposes of this study it was decided that this may have an impact on graduation and/or freshman grades. Therefore, a variable was created that reflected whether a student was enrolled in the Engineering Problem Solving when the new curriculum was implemented or before that time.

The indicator of economic status used in the study is *peer economic status (PES)*. PES, for the purpose of this study, is defined as the percent of students not enrolled in the Free Lunch at the student's high school during the four year period the student is expected to have attended high school. In other words, 100 percent minus the percentage of students at a high school enrolled in Free Lunch. Free Lunch is administered by the U.S. Department of Agriculture for students who come from a household which receives an income of less than 130% of the poverty guidelines. The poverty guidelines may change from year to year, but for example: the 2011 poverty guideline for a household with four members was \$22,350. So, if a family's income was less than \$29,055 then the students in that family were eligible for free lunch.

A combination of data retrieved from Louisiana Tech and information from the Common Core of Data available from the National Center for Education Statistics (NCES) was used to calculate PES for each student in the study. The institution's data matched students with high school codes. The Common Core of Data reports the percentage of students at each school who are enrolled in Free Lunch and it is connected to a high school code representing that specific high school. Using those high school codes, a crosswalk provided by Matthew Chingos and the Mellon Foundation [84], and the data provided by the Common Core of Data, PES was matched to students in the

Louisiana Tech University data. It should be noted that NCES adjusts data in order to protect the privacy of students; if a high school reports that all of the students are on Free Lunch, then data will indicate 95% of students are in the program so that no single student can be identified as eligible for free lunch. Again, Free Lunch participation of a high school is reported, but individual participation is not known. Using this information, the calculated PES is higher for students from high schools where few students are using the Free Lunch program and lower for students at schools with a larger percentage of students on Free Lunch.

Private schools are not as likely to enroll students in the Free Lunch program as public schools [85]. These schools are also not required to report student enrollment in the program, often resulting in a lack of data for students who attended these institutions [86]. Furthermore, only non-profit private (and public) schools are eligible for Free Lunch [87]. Therefore any other private schools, though they may receive funding through other avenues, do not receive assistance from the Free Lunch program. Of the participating private schools, only 8.2% of all K-12 students were approved for free or reduced-lunches [88]. Research has proven that participation in the program decreases as students become older [89] – indicating that a very small percentage of high school students attending non-profit private schools receive reduced lunches and an even smaller percent receive free lunches. From this information, it was decided that private school student's PES, if not reported, would be 100 percent – interpreted as zero percent of the students received free lunch.

PES has been used in multiple studies as a type of socioeconomic indicator, such as a study concerning graduation rates of engineering students [90]. Other studies have

used Free Lunch or Free and Reduced Lunch participation as measures of economic status as well [91, 89]. Studying K-12, researchers have shown that school-level and individual poverty measures are significant contributors to academic outcomes for students [92, 91, 93, 94]. Research has also provided evidence that *PES* may be more significant than a student's family economic status. The first piece of evidence is that peer expectations play a large role in postsecondary persistence [31, 95, 96] Secondly, schools are more likely to have fewer resources and lower parental participation at schools with a higher percentage of Free Lunch students [97, 98].

3.3.2 Transformation of Variables

All variables were assigned numeric values in order to analyze the data, and some required additional transformations. For example, high school GPA scores above 4.0 were converted to the 4.0 scale. Also, *high school rank* was transformed into a percentage by dividing the rank of the student by the number of students in that student's high school class, and then multiplying by one hundred. Next, that percentage was subtracted from 100 so that a higher percentage correlated with a higher standing in the class (as seen in Eq. 3.1).

Explicitly:

$$\text{High School Rank} = 100 - \left(\frac{\text{Rank in the Class}}{\text{Number in the Class}} \times 100 \right). \quad (3.1)$$

Sex was noted with a "0" for male students or a "1" for female students. *Race* had a similar transformation – "0" for White students and "1" for Black/African American students. The *state residency* variable also had two options: "0" for a non-resident and "1" for students from Louisiana. As mentioned earlier, *peer economic status (PES)* was a percentage based on the number of students at a high school enrolled in the Free Lunch

Program. A curriculum variable, abbreviated *LWTL*, was indicated with a “0” if a student was enrolled in the integrated freshmen engineering course and a “1” if a student was enrolled after the curriculum was changed to *Living with the Lab*. After cleaning the data and assigning numeric values to all variables, the resulting values and ranges of the variables are given in Table 1.

Table 1 Study One Predictor Variables and Their Ranges

Variable (Abbreviation)	Range/Values	
Sex (Sex)	0 = Male	1 = Female
Race (Race)	0 = White	1 = Black/African American
Louisiana Residency (State)	0 = Non-Resident	1 = Resident
Curriculum (LWTL)	0 = Integrated	1 = LWTL
High School Rank (HSRank)	3.0 – 100	
High School GPA (HSGPA)	1.6 – 4.0	
Peer Economic Status (PES)	0.54 – 100	
<i>ACT component scores</i>		
Science Score (ACT S)	7 – 36	
Mathematics Score (ACT M)	14 – 36	
English Score (ACT E)	11 – 36	
Reading Score (ACT R)	12 – 36	

3.3.3 Justification

All variables included in the study were chosen for multiple reasons. First, under the guiding theoretical framework of constructivism, the majority of the variables used were indicators of background or prior knowledge, such as *high school GPA* and *rank* as well as *ACT component scores*. Availability of data also played a role in which variables were chosen. Other factors have proven to be influential of freshmen grades, such as *self-efficacy* or being an introvert [48, 50, 51]; however, no measures of these factors were available to the researcher at the time of study.

Secondly, all of the variables were indicated in other research as influential to freshmen retention and grades. Multiple studies identified *high school GPA* as a predictor of success for students [99, 100, 101, 102]. Another study also identified a student's GPA as a positive indicator of retention to the second year along with state residency [103]. A separate study found that *non-Indigenous status* (for this study, whether or not a student identified themselves as being of Aboriginal and/or Torres Strait Islander descent) and *sex* (being female) were factors positively associated with first-year academic performance [104]. Another researcher found *sex* to be important in predicting freshman GPA; again being female was a positive factor [47]. However, other studies have indicated that *sex* is either not significant or a negative variable if a student is a female. For example, Honken and Ralston found that *sex* was not significant when studying retention to the second year [49]. Studying science, math, and engineering students, another set of researchers indicated that females were less likely to graduation [30].

Other studies have found that standardized test scores, referring to either SAT or ACT scores, are also positive indicators of success for college students [52, 101, 102, 105, 106, 107, 108]. For this study, ACT scores were used instead of SAT scores for two main reasons. First, 97 percent of Louisiana Tech University students admit ACT scores to the school while only three percent admit SAT scores [109]. Secondly, ACT has been almost as widely used as SAT in the past and as of 2010 more students took the ACT than the SAT [110].

Race and *economic status* were both significant factors in a study focusing on freshmen student retention at a university in the New York City area [111]. Lundy-

Wagner *et al.*, reviewing the literature of economic status, stated that “school context, and specifically high school poverty, is at least as important as, and possibly more important than, students’ individual sociodemographic characteristics on academic outcomes” [112].

3.4 Outcome Variables

3.4.1 Achievement Study

The studies involving grades included only those students who completed the specified class or withdrew from it; therefore, the observations were excluded if the student only audited the class or the grade was recorded as an incomplete. Withdrawing from the class was counted as an “F” in the class. At times, students withdraw from classes when they are in danger of failing the courses. Furthermore, if a student withdraws then he or she must take the course again before moving on to the next class which is the same outcome when a student fails the course. Additionally, other studies have grouped these cases similarly [29]. Using this information, class grades of “A,” “B,” “C,” “D,” and “F/W” became “4,” “3,” “2,” “1,” and “0” respectively.

3.4.2 Placement Study

Two different outcome variables were used for the second study. The first was an indicator of remediation, denoted by 1 if the individual took a remedial math class and 0 otherwise. This variable became an input variable for the second equation. The outcome variable for the second equation was *grade in Pre-Calculus*, denoted by either a “4,” “3,” “2,” “1,” or “0” accordingly.

3.4.3 Graduation Study

Graduation was defined as graduating from Louisiana Tech University within a six year time period and obtaining a degree in the student's chosen discipline as recorded in their freshmen year at the time the student was enrolled in UNIV 100. This binary variable was indicated with a 1 if the student did graduate in their major type within a six year time period and 0 if they did not.

3.5 Analysis Techniques

All of the studies in this research used a type of regression to analyze the data. Specifically, the first study utilized linear regression while the second study's method was regression discontinuity and the third study used logistic regression. An explanation of these regressions and their corresponding confidence intervals is given in the following pages. Unless otherwise noted, the statistic terminology used throughout this work will follow Montgomery, Peck, and Vining [113].

3.5.1 Linear Regression

To give a general explanation, multiple linear regression is fitting a line to a set of data which includes a dependent or outcome variable (y) and multiple independent or input variables (x). A mathematical representation of the regression is given Eq. 3.2:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon, \quad (3.2)$$

where each x represents the different independent variables, each β represents an unknown coefficient, ε is a random error component, and y is the dependent variable. A linear regression model is called an empirical model when the relationship between the dependent and independent variables is unknown and the model attempts to discover a

reasonable approximation of the unknown function. In other words, discover an estimate for each coefficient so that the equation produced gives a feasible estimate of the dependent variable given certain independent values. It is also possible to add interaction effects to a model, the combined effect of independent variables, which changes the equation slightly. The equation for a model with two independent variables and the interaction between the two variables would be as follows in Eq. 3.3:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \varepsilon, \quad (3.3)$$

where β_{12} is the coefficient for the interaction term and x_1x_2 is the interaction between x_1 and x_2 . If the hierarchical principle is applied, then an interaction term can only be added after each independent variable that makes up the interaction term already exists in the model. In choosing a regression model that fits the data, two major ideas should be considered. On one hand, you want a model that explains a large amount of variance and on the other hand, you want the simplest model possible. If the model explains a large amount of variance, it should be more accurate in predicting the outcome variable, but this usually requires more regressors. When regressors are added, it creates more noise in the model and wastes degrees of freedom. Therefore, the “best” model is a compromise between the simplest model and what explains the most variance. In order to discover a best model, three different methods are often used.

The first method is forward selection. Using this method, each independent variable is regressed against the outcome variable, and the variable which has the largest correlation with y is added to the model. The second step is to fit a linear regression model with two regressors: the variable added from step one and each of the other variables available in the model. Again, whichever pair of variables has the largest

partial correlations to the response variable are kept for step three. This process continues until none of the partial correlations are less than a predefined cutoff correlation value.

A second commonly used method is backward elimination; it is the reverse of forward selection. Instead of starting the regression with only one regressor, all the regressors are initially included in the model. Each step deletes a variable that has the smallest partial correlation until the variables with the smallest partial correlation is greater than a predefined cutoff correlation value.

The last method is a mix between forward selection and backward elimination; it is called stepwise selection. The initial model includes no variables and the first step adds a variable to the model. However, after each step in which a variable is added, the model also checks that variables previously added are still important to the model given that other variables are already in the model. Basically, the algorithm checks that variables previously added are not redundant. The process ends when the partial correlation of a new variable is either greater than the predefined “take out” cutoff or less than the predefined “keep in” cutoff.

Besides considering the correlation between an independent variable and the response variable, it is also possible to use other criteria to judge if a variable is added to the model. Two other options are most often used: the coefficient of multiple determination (R^2) and adjusted R^2 [113].

The coefficient of multiple determination is a function of the total sum of squares of a regression model and the residual sum of squares. The equation for this coefficient could be given as seen in Eq. 3.4:

$$R^2 = 1 - \frac{SS_{Res}}{SS_T}, \quad (3.4)$$

where SS_{Res} is residual sum of squares and SS_T is the total sum of squares. When using this criteria to choose a model, the algorithm searches for the simplest model with the largest R^2 . In other words, the model with the smallest number of regressors where R^2 is not significantly different from the previous model that was tested. Using this coefficient does have a downfall however – the largest value of R^2 will always happen when all the regressors are in the model. In other words, the value will always decrease when variables are removed from the model. To compensate for this issue, a different coefficient can be used: adjusted R^2 .

Adjusted R^2 is similar to the coefficient of multiple determination except that it also considers the number of regressors in the model and the size of the data. The equation could be written as seen in Eq. 3.5:

$$R_a^2 = 1 - \frac{(n-1)}{(n-p)}(1 - R^2), \quad (3.5)$$

where n is sample size and p is the number of regressors in the model including the intercept. As with the coefficient of multiple determination, using the adjusted R^2 as the criterion for choosing a model means searching for the simplest model with largest adjusted R^2 .

For the purpose of this research, to choose a best model when using linear regression, two aspects are considered: the significance of the variables in the model and the change in adjusted R^2 . All variables in the model must be significant. Also, adding a variable to the model must increase the adjusted R^2 by at least five-tenths of a percentage point. There is no common rule of thumb for this measure; the researcher chose this policy.

3.5.2 Regression Discontinuity

The regression discontinuity (RD) design was originally created by Thistlewaite and Campbell for research in psychology and education [114]. In essence, it is a pre-post, comparison group design. For studies analyzing remedial effects or the effectiveness of placement policies, the pre-test is a score (also referred to as the assignment variable) that is used to place students in a specific course. The administered program is the remedial course; one group of students is enrolled in the program (treatment group) and the other is not (control group). The post-test score is usually a measure of achievement in the course which all students were enrolled, such as final grade in the class. This RD design is recognized as the only quasi-experimental design which fulfills the requirements for establishing a causal relationship [115].

Therefore, regression discontinuity takes advantage of data that is separated by a specific factor, such as a cutoff score dictating whether or not a student takes a remedial class [116]. This is one of four specific requirements that must be met in order to use this type of analysis – a discontinuous jump must be present in the data [117]. The second requirement is that the assignment variable is not caused by the treatment [118]. Third, the participants must not be able to manipulate treatment [117]. However, if a large enough percentage of students (more than five percent) are no-shows (assigned to the treatment group by the assignment variable, but not treated) or crossovers (assigned to the control group by the assignment variable, but receive treatment) then a *sharp* regression discontinuity is no longer possible and instead *fuzzy* regression discontinuity must be used [119]. The probability of being assigned to a treatment condition jumps from 0 to 1 while the treatment assignment becomes fuzzy if assignment is influenced by

factors other than, say, a cutoff score. Therefore, applying a sharp RD design when treatment assignment is fuzzy can produce bias in estimates.

In order to provide valid causal evidence, the fourth standard is that all students just above and just below the jump must share similar relevant characteristics [117]. Levin and Calcagno state that only students who share similar academic preparedness and backgrounds should be compared [120]. After the four requirements to use RD are met, the next step is to implement the analysis. For fuzzy regression discontinuity, there are two stages of equations. The first equation (Eq. 3.6) is as follows:

$$T = \alpha_0 + \alpha_1 D + \alpha_2 r + f(r) + \varepsilon_1. \quad (3.6)$$

The definition of each variable is given in Table 2.

Table 2 Definitions of Variables for First Stage Equation

Variable	Definition
T	1 if an individual enrolled in a remedial math class, and 0 otherwise
D	1 if an individual is assigned to take a remedial math class based on the cutoff rule, and 0 otherwise
r	transformed ACT cutoff score centered around the cutoff point ($r = \text{student's ACT math score} - 26$)
$f(r)$	relationship between r , the transformed cutoff score, and D , the assignment to remediation
ε_1	random error for first equation, assumed to be identically and independently distributed

The relationship between r and D can be specified in multiple ways. The most complicated option is defining $f(r)$ is seen in Eq. 3.7:

$$r * D + r^2 + r^2 * D + r^3 + r^3 * D. \quad (3.7)$$

According to Shadish, Cook, and Campbell, initial regression discontinuity models should initially contain all these variables (interaction and higher order ones) in

order to test if any are significant [121]. Arguing the case to include higher order terms from a different perspective, Campbell and Russ noted that underestimating a model “can lead to pseudo-effects, whereby the reverse error, overfitting...should not” [122].

Initially, all of the terms included in $f(r)$ are included in the model. After the initial model is analyzed, the least significant term is deleted. This process continues until only significant terms are left in the model and the model is significant. Therefore, if none of the interaction or higher terms are significant to the model, then none are included.

After determining final model of the first stage equation (which is used to obtain the value for T), the second equation is seen in Eq. 3.8:

$$Y = \beta_0 + \beta_1 T + \beta_2 r + f_2(r) + \varepsilon_2, \quad (3.8)$$

where Y is the outcome variable (*grade in Pre-Calculus*) and $f_2(r)$ = the relationship between T and r .

Like the first stage equation, the relationship between T and r can take multiple forms by possibly including higher order and interaction terms. It is also possible that no interaction or higher order terms are included in the model if none of them are significant. Again, the most complex option for the relationship between the two variables is tested first. The equation is as follows in Eq. 3.9:

$$f_2(r) = r * T + r^2 + r^2 * T + r^3 + r^3 * T. \quad (3.9)$$

Like the first stage equation process, if a variable is not significant to the model, then it is deleted. This process continues until all variables in the model are significant and the model itself is significant. After an analysis is completed, a significant coefficient for T suggests that there is a significant discontinuity at the cutoff score- there

is a main effect of remediation [123]. If the placement coefficient (r) is negative, it is an indication that the cutoff scores should potentially be lowered [124, 125]. On the other hand, a large positive coefficient indicates cutoff scores may need to be raised. Another option is a placement coefficient that is close to zero or modestly positive which indicates that the placement policy is well-designed. In other words, students just above and below the cutoff are doing similarly well.

3.5.3 Logistic Regression

An alternative to linear regression is logistic regression. This regression can only be used when the dependent or response variable is categorical. Most often, the values of the response variables are either 0 or 1 and represent failure or success. A model for logistic regression takes the form seen in Eq. 3.10:

$$\ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \cdots + \beta_kx_k + \varepsilon, \quad (3.10)$$

where π is the probability that the response variable is a success (usually indicated with a “1” in the data) and $1 - \pi$ is the probability of a failure. Therefore, logistic regression predicts the *probability* of a certain outcome. Often accompanying logistic regression, odds ratio (OR) is a value derived from the model which represents the change in the probability of success when a certain regressor, x , is increased by one unit. To compute the odds ratio estimate, Eq. 3.11 is used:

$$\left(\frac{odds_{x_i+1}}{odds_{x_i}}\right) = e^{\beta}, \quad (3.11)$$

where β is the coefficient estimate for regressor x . As seen in the formula, odds ratio is a function of odds. A contingency table is helpful in explaining this. For example, a group

of people are given either a new drug or a placebo, and some of the participants become ill and some do not. This situation is shown in Table 3.

Table 3 Ill and Healthy Patients Given New Drug or Placebo

	Ill	Healthy
New Drug	a	c
Placebo	b	d

The odds of a participant that is given the placebo becomes ill is the probability of becoming ill given that the participant is on the placebo divided by the probability of staying healthy given that the participant is on the placebo, or in equation form b/d . The odds of a participant being given the drug becoming ill is the probability of becoming ill given that the participant is on the new drug divided by the probability of staying healthy given that the participant is on the new drug. The equation for this would be a/c . Therefore, to compute the odds ratio estimate of a participant being given the drug becoming ill the equation would be the one seen in Eq. 3.12:

$$OR = \frac{a/c}{b/d} = \frac{ad}{bc}, \quad (3.12)$$

Interpretation of the odds ratio is fairly simple. Let the OR be 2.3. Then the odds of a participant on the new drug becoming ill is 2.3 times that of a participant on the placebo [113]. From this information, it is apparent that an OR close to 1 has little impact while an OR farther from 1 is more impactful. Accordingly, if the confidence interval for the OR of a regressor includes 1 then it is not significant.

Another common measure associated with logistic regression is the Hosmer-Lemeshow (HL) statistic which is used to measure goodness of fit. This statistic

asymptotically follows a chi-square distribution with $g - 2$ degrees of freedom where g is the number of groups into which the data is split. Like a Pearson chi-square statistic comparing observed and expected frequencies, the formula for HL statistic is seen in Eq. 3.13:

$$HL = \sum_{j=1}^n \frac{(O_j - N_j \pi_j)^2}{N_j \pi_j (1 - \pi_j)}. \quad (3.13)$$

The definitions for the variables of Equation 3.13 are given in Table 4.

Table 4 Definitions of Variables for HL Statistic Formula

Variable	Definition
O_j	number of observed successes
N_j	number of observations in the j th group
π_j	average estimate success probability in the j th group

Smaller values of this statistic (accompanied by large p -values) are better as large values generally indicate the model is not an adequate fit. A second method of deciding adequacy through the HL statistic is dividing the statistic by the degree of freedom of a model ($g - f$) where f the number of independent variables in the model [113]. If this ratio is close to unity, the model is deemed adequate. Hosmer and Lemeshow state that at least six groups should be used and ten is the suggested (and the most common) number of groups to use [126, 127]. For this study, ten groups were used to calculate the statistic.

For this research, to choose a best model using logistic regression, two aspects are considered. First, all variables in the model should be significant. Secondly, the HL statistic must indicate that the model is a good fit for the data.

3.5.4 Confidence Intervals and Significance

In both linear and logistic regression, confidence intervals can be calculated for each coefficient estimate [113]. A significance level, denoted by alpha (α), is necessary to compute the interval. For the purpose of this research, all studies use an alpha level of 0.05. In linear regression, a $100 * (1 - \alpha)$ confidence interval for each coefficient estimate can be determined by Eq. 3.14:

$$\beta - t_{\frac{\alpha}{2}, n-2} se(\beta) \leq \beta \leq \beta + t_{\frac{\alpha}{2}, n-2} se(\beta), \quad (3.14)$$

The definition of each variable seen in the confidence interval equation are given in Table 5.

Table 5 Definitions of Variables for Coefficient Confidence Interval

Variable	Definition
β	coefficient estimate for a certain regressor
$se(\beta)$	standard error
$t_{\frac{\alpha}{2}, n-2}$	student's t distribution with $n - 2$ degrees of freedom

This confidence interval equation returns an interval for the coefficient estimate, and if using an alpha level of 0.05 then there is a 95% chance that the true value of the coefficient is in the given range. This is also true for confidence intervals determined for the logistic regression coefficient estimate and odds ratio estimate. With logistic regression, the formula for the confidence interval is very similar (as seen in Eq. 3.15):

$$\beta - Z_{\frac{\alpha}{2}} se(\beta) \leq \beta \leq \beta + Z_{\frac{\alpha}{2}} se(\beta), \quad (3.15)$$

Table 6 contains the definitions for each variable from Equation 3.15.

Table 6 Definitions of Variables for Coefficient Confidence Interval: Logistic Regression

Variable	Definition
β	coefficient estimate for a certain regressor
$se(\beta)$	standard error of that coefficient
$Z_{\frac{\alpha}{2}}$	standard normal distribution

It is also possible to calculate a confidence interval for the odds ratio estimate. As previously mentioned, the odds ratio and the accompanying confidence interval can help determine the significance of a regressor. The general structure for the OR confidence interval is given in Eq. 3.16:

$$Estimate \pm Confidence\ Coefficient * Standard\ Error. \quad (3.16)$$

Another indicator of significance is the p-value associated with each regressor. In general, a smaller p-value indicates more significance than a larger value. If an alpha level of .05 is chosen, then any coefficient estimate and variable that produce a p-value of less than .05 is considered significant. If the value is less than .01, then it is more significant, and if the p-value is less than .001 then it is very significant.

3.6 Protection of Human Subjects

The research project began after a research proposal was approved by the doctoral student's (researcher) advisor along with the dissertation committee and permission was gained from the Louisiana Tech Institutional Review Board to conduct the study. The majority of the data were retrieved from a single source - Louisiana Tech University's academic records. The student data were de-identified before the researcher collected it. Re-identification of the data required a securely stored and password protected key only available to the advisor and the key was required for all data updates.

Therefore, no individually identifiable information was disclosed. A waiver of written consent from all participants was given as the research met these four requirements as stated under 45 CFR 46.116(d):

- a) the research involves no more than minimal risk to the subjects;
- b) the waiver or alteration will not adversely affect the rights and welfare of the subjects;
- c) the research could not practicably be carried out without the waiver or alteration; and
- d) whenever appropriate, the subjects will be provided with additional pertinent information after participation.

This research meets the above requirements as follows:

- a) By using information that is currently being collected on each participant for institutional records, this research poses no risks to the participants beyond the risks normally associated with research conducted in the Office of Institutional Research at Louisiana Tech University.
- b) The justification for collecting and analyzing this type of data can be found in 20 USC §1232g(b)(1)(F);

(b) Release of education records; parental consent requirement; exceptions; compliance with judicial orders and subpoenas; audit and evaluation of Federally-supported education programs; recordkeeping.

(1) No funds shall be made available under any applicable program to any educational agency or institution which has a policy or practice of permitting the release of educational records (or personally identifiable information contained therein other

than directory information, as defined in paragraph (5) of subsection (a)) of students without the written consent of their parents to any individual, agency, or organization, other than to the following—

(F) organizations conducting studies for, or on behalf of, educational agencies or institutions for the purpose of developing, validating, or administering predictive tests, administering student aid programs, and improving instruction, if such studies are conducted in such a manner as will not permit the personal identification of students and their parents by persons other than representatives of such organizations and such information will be destroyed when no longer needed for the purpose for which it is conducted.

The research goal was to improve how students learn by discovering why students succeed or fail as they maneuver through the engineering curriculum at the university.

- c) This research would be impractical without the waiver of consent. Tracking down every student that has attended Louisiana Tech since 1990 would be cost prohibitive.
- d) The research is available to students through journal articles, conference proceedings, and by email request.

Information concerning peer economic status was taken from the National Center for Education Statistics (NCES) Common Core of Data. Again, no individual's identifying information was available to the researcher.

CHAPTER 4

STUDY ONE: ACHIEVEMENT

4.1 Approach and Justification

For the first study, multiple linear regression was used to determine variables that were significant for engineering freshmen in terms of first quarter grades in Pre-Calculus and a freshman engineering course (Engineering Problem Solving I). Other studies have implemented the same method when testing similar hypotheses. For example, a study modeling college success in terms of first semester GPA of college students also used multiple linear regression to determine predictive factors [51]. Also using a multiple linear regression, a different set of researchers predicted student's first-year cumulative GPA using ten potential variables derived from demographic and high school information [47]. In addition, Brown, Halpin, and Halpin analyzed the effect of high school mathematics preparation on pre-engineering GPA using regression analysis as well [128]. Using step-wise and best subset linear regression, Veenstra, Dey, and Herrin studied factors predicting academic success (first-year GPA) for both engineering and non-engineering freshmen students [52]. A study involving minority engineering students at the University of Akron used regression analysis to determine an equation for predicting undergraduate GPA in terms of *ACT score* and *high school GPA* [129].

4.2 Predictor and Outcome Variables

The predictor variables for the first study were: *high school GPA, high school rank, sex, race, ACT math score, ACT English score, ACT reading score, ACT science score, state residency, peer economic status, and enrollment in Living with the Lab*. Initially, 2989 students were identified as freshman engineering majors enrolling between 2006 and 2014. Of that group, 185 students were removed because of a missing ACT math score along with 161 with missing high school rank and 7 missing high school GPA. As previously mentioned, multiple categories of race had small sample sizes and were removed accordingly; the total number of students removed was 283. The last two steps in cleaning the data checked for missing peer economic status (PES) or grade in Pre-Calculus; 118 more students were deleted for a final sample size of 2235 students. Similar steps were taken to clean the data used for the engineering course regression model with similar results; the final sample size consisted of 2204 students for that group.

The outcome variables for study one were (1) *grade in Pre-Calculus* for the first model and (2) *grade in the first freshmen engineering class* (Engineering Problems Solving I) for the second model.

4.3 Participants

The participants for the first study were first-time-in-college (FTIC) freshmen engineering majors enrolled at Louisiana Tech University from 2006 to 2014. Possible majors of the students during the time span of the study are: basic, biomedical, civil, chemical, cyber, electrical, industrial, mechanical, and nanosystems engineering.

This study specifically analyzed engineering freshmen's Pre-Calculus and Engineering Problem Solving I grades using two different regression models. After removing observations with missing data, the Pre-Calculus sample included 2235 students and the freshmen engineering class sample included 2204 students (largely a subset of the Pre-Calculus sample as most students take Pre-Calculus and the engineering class concurrently).

The final sample included 87.9% White and 12.1% Black students. As for sex, about sixteen percent of the students were female and eighty-four percent were male – an unbalanced sample, but not out of place for a group of undergraduate engineering majors. Over 90% were students from Louisiana. These descriptive statistics are listed in Table 7.

Table 7 Study One Sample by Race, Sex, and State

N	White	Black	Female	Male	In State
2235	87.9%	12.1%	15.9%	84.1%	90.1%

4.4 Method

For the achievement study (study one), linear regression with forward selection was implemented in order to predict final grade in Pre-Calculus and Engineering Problem Solving I. Significance of variables and increase in adjusted R^2 were used to determine the best models. Backward elimination and stepwise regression were also implemented to check forward selection results.

4.5 Expected Outcomes

As a student's ACT math score is used to place students in initial math classes at Louisiana Tech University, it was expected that this variable would be significant in terms of freshmen math and engineering grades. Additionally, other studies have found standardized test scores or prior math background are influential to student's freshmen GPA [11, 40, 45, 52, 44]. However, based on previous findings, it was also expected that other variables would prove to be significant influences (positive or negative) of the grades, such as *high school GPA*, *high school rank*, and *sex* according to previous research [29, 30, 31, 32, 44, 45, 52].

4.6 Limitations

The first limitation is that the data analyzed for study one came from a single university instead of multiple institutions. Including more data from different universities would have given more validity to the results and increased the generalizability of the study. On the other hand, it allowed a detailed examination of a particular context that would be clouded by institutional variation. One research study stated the importance of studying students at a specific institution this way, "Engineering colleges must understand their student population in order to design interventions that will improve retention of their students" [49].

A second shortcoming was that due to small sample sizes, only two races were included in the study – White and Black. Other races/ethnicities, such as Hispanic, Asian, or Pacific Islander were not included as they collectively represented a very small percentage of the total population of participants. Including these small cell sizes could have resulted in overgeneralization, insufficient statistical power, and compromised

anonymity. Furthermore, the data did not contain variables such as *marital status*, *self-efficacy*, and *transfer credit/dual enrollment*. Other studies have indicated that these variables may have an effect on first-year grades of freshmen students and even graduating with an engineering degree [12, 39, 130, 131]. The inclusion of said variables could have potentially changed the outcomes of the analyses. Unfortunately, measures of these factors were unavailable to the researcher.

Innate shortcomings of *PES* were also a limitation. Researchers have indicated that there are systematic biases in the variable, such as a decline in participation of the Free Lunch as students became older and that the program is based on outdated poverty guidelines [89, 132]. Additionally, private schools which did not report participation in the program were considered as having zero percent free lunch; while this is a reasonable assumption, the author recognizes that this is an imperfect measure.

The final grade of “F” and “W” (denoting if a student withdrew from the class) were grouped together for this study. Though either of these final results ensures that a student must retake the class in order to continue to the next mathematics or engineering class, it is possible that a student withdraws for reasons other than being at danger of failing the class.

Another limitation of the research lies in the type of method used – observations that contained a missing variable were not able to be processed and therefore were removed from the study. This allowed only a subset of student information to be analyzed.

As already mentioned, no transfer students were included in the studies as these students are not required to take UNIV 100; however, all first-time freshmen are required

to take the class and therefore enrollment in this class was used to identify engineering freshmen. Another group of students were not included – those who bypassed Pre-Calculus and enrolled in a higher level math class. Starting in the summer of 2011, students with appropriate ACT scores were given the opportunity to take an exam. If the students passed, they were allowed to enroll in a higher level math course. It should be noted, however, that the number of students who pass the exam is relatively low – the average is five students per year.

Another set of students not included in the study are those that have taken AP Calculus in high school and passed the appropriate exams. For example, students who take the AP Calculus AB exam and score a 4 or 5 received credit for both Pre-Calculus and Calculus I. Those who take the BC Exam and score a 3, 4, or 5 also receive credit for Pre-Calculus and Calculus.

4.7 Results for Study One: Achievement

Before the linear regression models for Pre-Calculus and Engineering Problem Solving I were completed, a correlation matrix was analyzed to ensure that no two variables were highly correlated. When two variables are highly correlated, which can vary in definition but for the purpose of this work a correlation coefficient above 0.8 will be considered high, there are possible repercussions – such as a change in the sign of a coefficient or a large shift in the significance of a variable [133, 134]. Although correlation between variables is not the only possible cause of a sign change, a change in sign can indicate the presence of correlation [134]. If two variables are highly correlated then the results of a regression may be skewed. One of the options to deal with high correlation is to drop one of the variables [133].

Analyzing the variables used in the study, four sets of variables returned correlation coefficients above 0.60. These are listed in Table 8. Moderately high correlations existed between different ACT component scores, but were kept in the model as some correlation between the scores was expected and the correlations were not extreme. A significantly high correlation, also seen in Table 8, was discovered between two other variables: *high school rank* and *high school GPA*.

Table 8 Correlation Coefficients above 0.6

HSGPA/HSRank	ACT		
	S/M	S/R	R/E
0.83	0.60	0.61	0.64

To verify that the variables were correlated as well as decide which variable to drop from the study, another step was taken. Three linear regression models were created where the outcome variable was grade in Pre-Calculus and the input variables were *high school GPA* and *high school rank* (each individually regressed against grade and the third model including both variables regressed against the grade). From these models, it was possible to observe if a change in sign occurred from the individual regression models to the combined model where both variables were present. As seen in Table 9, both *high school GPA* and *rank* were positive and significant predictors of *grade in Pre-Calculus* when regressed individually. The third model in the table showed that in the model with both variables the sign of the coefficient for *high school rank* changed from positive to negative which lent further proof that the two variables were highly correlated.

Table 9 Linear Regressions with High School GPA and Rank Against Grade in Pre-Calculus

Model	Coefficient	Std. Error	Conf. Interval		Adjusted R ²	p-value
			2.5%	97.5%		
Model 1					22.8%	
HSGPA	1.7002	0.0662	1.5704	1.8301		< 2e-16 ***
Model 2					14.3%	
HSRANK	0.0253	0.0013	0.0228	0.0279		< 2e-16 ***
Model 3					22.8%	
HSGPA	1.8584	0.1186	1.6257	2.0910		< 2e-16 ***
HSRANK	-0.0036	0.0022	-0.0079	0.0008		0.108

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

Notice that the first model with only *high school GPA* produced a larger R² than the model only involving *high school rank*. For the model including both variables, the R² value did not increase from Model 1 and *high school GPA* was significant while *high school rank* was not (which is another possible sign of high correlation). Using the outcome variable of *grade in the engineering course* produced results consistent with the models for grade in Pre-Calculus. The *high school GPA* model explained more variance than the *high school rank* model and the sign of the coefficient as well as significance of *high school rank* changed when both variables were included in the model.

With these results, it was determined that *high school rank* should not be included in the study as *high school GPA* explained more variance. Additionally, *high school GPA* has proven to be a significant predictor of freshmen grades in multiple studies, including a preliminary study analyzing engineering students at Louisiana Tech University [135].

The first research question for this study asked, “*Is ACT math score a significant influence on the final grades of Pre-Calculus for engineering students?*” In order to answer this, a linear regression model testing the significance of *ACT math score* against

Pre-Calculus grades and then engineering grades was implemented, respectively (shown in Table 10).

Table 10 ACT Math as a Predictor of Pre-Calculus and Engineering Grade

Model	Coefficient	Std. Error	Conf. Interval		Adjusted R ²	p-value
			2.5%	97.5%		
<i>Pre-Calculus</i>					13.6%	
ACT M	.138	.007	.123	.152		< 2e-16 ***
<i>Engineering</i>					10.2%	
ACT M	.123	.008	.107	.138		< 2e-16 ***

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

As seen, *ACT math score* was significant for both models. Therefore, *ACT math score* is a significant influence on *grade in Pre-Calculus* and *grade in Engineering Problem Solving* for engineering students.

4.7.1 Pre-Calculus Models

The second research question asked if other variables, besides *ACT math score*, also had a significant influence on grades for engineering students in addition to *ACT math score*. In order to test this, a linear regression model with forward selection was employed for two models; one with an outcome of *grade in Pre-Calculus* and one for *grade in the engineering class*. The regressors for the models included: *high school GPA*, *state residency*, *sex*, *race*, *PES*, *enrollment in either the integrated engineering curriculum or LWTL*, *ACT English score*, *ACT reading score*, and *ACT science score*. Notice that *high school rank* was not included as it was highly correlated to *high school GPA*.

Before running the analysis, a decision also had to be made concerning the inclusion of interaction terms in the model. The researcher chose to use the hierarchical principle – an interaction term can only be added if both of the variables that make up the interaction are included in the model. When interactions are included without the main effects, the meaning can be changed; the interaction terms then also contain the main effect terms.

For the first linear regression model, the outcome variable was *grade in Pre-Calculus*. To consider a new model better than the previous one, the new model must have increased the amount of variance explained by at least five-tenths of a percentage and all variables in the model should be significant. Using forward selection to add the variable that explains the most variance in the model (using adjusted R^2), the first variable included was *high school GPA*. It accounted for 22.8% of the variance in the model and was significant. The second variable added to the model was *ACT math score*. Adding this term increased the adjusted R^2 to 27.3% and both variables in the model were highly significant. The third variable that was suggested to add to the model was *PES*. This resulted in an increase of over a percentage point in the adjusted R^2 and all variables in the model were significant. The fourth term included in the next model was *sex*. After adding this term the adjusted R^2 increased more than half a percent. The curriculum variable, *LWTL*, was the fifth variable suggested to add to the model; it added five-tenths of a percentage point to the amount of variance explained. *ACT reading score* was added next. At this point the amount of variance explained only increased by two-tenths. The results of these models are detailed in Table 11.

Table 11 Pre-Calculus Models

Model	Coefficient	Std. Error	Conf. Interval		Adjusted R ²	p-value
			2.5%	97.5%		
Pre-Calculus Model 1					22.8%	
HSGPA	1.70	0.066	1.61	1.89		< 2e-16 ***
Pre-Calculus Model 2					27.3%	
HSGPA	1.41	0.069	1.32	1.62		< 2e-16 ***
ACT M	0.085	0.007	0.068	0.099		< 2e-16 ***
Pre-Calculus Model 3					28.4%	
HSGPA	1.47	0.069	1.336	1.607		< 2e-16 ***
ACT M	0.078	0.007	0.063	0.092		< 2e-16 ***
PES	0.009	0.001	0.006	0.011		4.36e-09 ***
Pre-Calculus Model 4					29.0%	
HSGPA	1.41	0.070	1.277	1.551		< 2e-16 ***
ACT M	0.081	0.007	0.067	0.095		< 2e-16 ***
PES	0.009	0.001	0.006	0.011		4.17e-09 ***
SEX	0.303	0.069	0.168	0.439		1.18e-05 ***
Pre-Calculus Model 5					29.5%	
HSGPA	1.41	0.070	1.271	1.545		< 2e-16 ***
ACT M	0.082	0.007	0.067	0.096		< 2e-16 ***
PES	0.009	0.001	0.006	0.011		4.27e-09 ***
SEX	0.319	0.069	0.183	0.454		4.00e-06 ***
LWTL	-0.224	0.054	-0.224	-0.119		3.10e-05 ***
Pre-Calculus Model 6					29.7%	
HSGPA	1.43	0.070	1.292	1.567		< 2e-16 ***
ACT M	0.091	0.008	0.076	0.107		< 2e-16 ***
PES	0.009	0.001	0.006	0.012		2.33e-09 ***
SEX	0.333	0.069	0.198	0.468		1.47e-06 ***
LWTL	-0.218	0.054	-0.323	-0.113		4.67e-05 ***
ACT R	-0.017	0.006	-0.028	-0.005		0.00379 **

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

As seen from observing the results of Model 5 and Model 6, the adjusted R² did not increase significantly by adding the term *ACT reading score*. Also, the sign of the coefficient for *ACT reading score* was negative; if regressed against grade by itself, the coefficient for that variable was positive. A possible explanation of this is that *ACT math score* and *ACT reading score* were moderately correlated and this relationship affected the model. Regardless of the reason for the change in sign, given that the amount of

variance explained by the last model only increased slightly, Model 5 was chosen as the final model before considering interaction terms. Possible interaction terms to add to the model would be any of the ten interactions of the remaining variables. Again using forward regression, further analysis indicated that the inclusion of an interaction term would increase the amount of variance explained by the model, but at the cost of loss of significance of other variables (Table 12). The first new term added to the model was the interaction between *high school GPA* and *ACT math score*.

Table 12 Pre-Calculus Model with Interaction Term

Model	Coefficient	Std. Error	Conf. Interval		Adjusted R ²	p-value
			2.5%	97.5%		
<i>Pre-Calculus Model 7</i>					30.3%	
HSGPA	-0.923	0.439	-1.783	-0.062		0.0357 *
ACT M	-0.244	0.061	-0.364	-0.125		6.48e-05 ***
PES	0.009	0.001	0.006	0.011		2.79e-09 ***
SEX	0.318	0.068	0.183	0.452		3.75e-06 ***
LWTL	-0.237	0.053	-0.341	-0.132		9.35e-06 ***
HSGPA*ACT M	0.092	0.017	0.058	0.126		8.45e-08 ***

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

When the interaction term between *high school GPA* and *ACT math score* was added to the model, the explanation of variance increased by eight-tenths of a percentage from Model 5. However, it should be noted that the interaction affected the signs of the coefficients for *high school GPA* and *ACT math score*; additionally, *high school GPA* was no longer as significant to the model. It was decided that this relatively small increase in explanation of variance was not worth the added complexity or changes in the model. Therefore the chosen model for predicting calculus grade remained Model 5. This

model included *high school GPA*, *ACT math score*, *PES*, *sex*, and *Living with the Lab enrollment*. The resulting equation from the results of Model 5, Eq. 4.1, was as follows:

$$\begin{aligned} \text{Predicted Grade in Pre-Calculus} = & \quad (4.1) \\ & -5.59 + 1.41(\text{HSGPA}) + 0.082(\text{ACT M}) + 0.009(\text{PES}) \\ & + 0.319(\text{SEX}) - 0.224(\text{LWTL}) \end{aligned}$$

To confirm these results, linear regression with backward elimination was also implemented. Using this method the first variable to be eliminated was *ACT science score* followed by *state residency*. At this point in the model, all variables left were indicated as significant (Table 13).

Table 13 Pre-Calculus Model with Backward Elimination: All Variables Significant

Model	Coefficient	Std. Error	Adjusted R ²	p-value
<i>Pre-Calculus Model 8</i>			29.9%	
HSGPA	1.42	0.072		< 2e-16 ***
ACT M	0.089	0.009		< 2e-16 ***
PES	0.009	0.001		4.5e-10 ***
SEX	0.304	0.070		1.34e-05 ***
LWTL	-0.223	0.054		3.16e-05 ***
ACT R	-0.023	0.007		0.000709 ***
RACE	0.185	0.083		0.024975 *
ACT E	0.016	0.008		0.049299 *

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

Though these eight variables were all significant, not all of the variables added a large amount of variance explanation to the model. In alignment with the requirements used in forward selection, a model with more variables had to explain at least five-tenths of a percentage point more variance than a model with less variables to be considered the better model. As removing *ACT English score* and *race* respectively decreased the adjusted R² by less than two-tenths of a percentage, each of the terms were removed. In

other words, both variables added two-tenths of a percentage or less to the adjusted R^2 value. Removing *ACT reading score* from the model resulted in an adjusted R^2 value of 29.5 percent. A model excluding the *LWTL* variable produced an adjusted R^2 of 29 percent. These results can be seen in Table 14.

Table 14 Pre-Calculus Models with Backward Elimination: Models 9 – 12

Model	Coefficient	Std. Error	Adjusted R^2	p-value
<i>Pre-Calculus Model 9</i>			29.8%	
HSGPA	1.45	0.071		< 2e-16 ***
ACT M	0.095	0.008		< 2e-16 ***
PES	0.009	0.001		2.58e-10 ***
SEX	0.317	0.069		5.11e-06 ***
LWTL	-0.219	0.054		4.33e-05 ***
ACT R	-0.016	0.006		0.00554 **
RACE	0.187	0.083		0.02374 *
<i>Pre-Calculus Model 10</i>			29.7%	
HSGPA	1.429	0.070		< 2e-16 ***
ACT M	0.091	0.008		< 2e-16 ***
PES	0.009	0.001		2.33e-09 ***
SEX	0.333	0.069		1.47e-06 ***
LWTL	-0.218	0.054		4.67e-05 ***
ACT R	-0.017	0.006		0.00379 **
<i>Pre-Calculus Model 11</i>			29.5%	
HSGPA	1.408	0.070		< 2e-16 ***
ACT M	0.082	0.007		< 2e-16 ***
PES	0.009	0.001		2.33e-09 ***
SEX	0.319	0.069		1.47e-06 ***
LWTL	-0.224	0.054		4.67e-05 ***
<i>Pre-Calculus Model 12</i>			29.0%	
HSGPA	1.41	0.070		< 2e-16 ***
ACT M	0.081	0.007		< 2e-16 ***
PES	0.009	0.001		4.17e-09 ***
SEX	0.303	0.069		1.18e-05 ***

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

Notice that the final three models were seen previously in the forward selection process and the results are the same. Therefore, backward elimination also indicated that

a model including *high school GPA*, *ACT math score*, *PES*, *sex*, and *enrollment in LWTL* was the best. Stepwise regression was also implemented, and the results concurred with the final models from forward selection and backward elimination.

4.7.2 Discussion of Pre-Calculus Results

According to the regression analysis, five variables were significant in relation to *grade in Pre-Calculus*. The amount of variance explained by this model was 29.5%. This result is similar to other studies. A study by Astin, Korn, and Green concerning the retention and satisfaction had results explaining 12% of the variance [105]. Tross *et al.* created models to predict college performance and retention resulting in models which explained 29% of the variance [136]. Another work predicting *cumulative GPA* using *SAT score*, *high school rank* and *gender* resulted in models explaining 18% of the variance [11].

As for application of these results, some possible avenues are: a recruitment tool, retention, and freshmen math placement requirements. For recruiters, the analysis suggested that students with both a high ACT math score and high school GPA would do well in Pre-Calculus. The final model for predicting *grade in Pre-Calculus* also indicated that a higher PES score has an influence on this outcome; while this variable is not an applicable one for recruiting, it could be used to help retain students. According to the model, a lower PES score means a student is less likely to pass Pre-Calculus, especially if the student also has a high school GPA and/or ACT math score on the lower end of the spectrum. The implications are similar in regards to sex- a male student with a low high school GPA, ACT math score and PES may be at risk of failing Pre-Calculus. The engineering program, perhaps through a student success specialist, can be aware of these

facts and be prepared to assist students in this situation. Though the model indicated that *Living with the Lab enrollment* is influential to *grade in Pre-Calculus*, more research is needed to determine whether this new program contributed positively to the success of an engineering student.

Moving on to placement, Louisiana Tech University primarily uses students' ACT math scores to place students in the first math class. However, the regression model indicated that *high school GPA* plays the largest role in determining the final *grade in Pre-Calculus* (*HSGPA* alone explained 22.7% of the variance while *ACT math score* alone explained only 13.6% and the other variables explained even less). With these results, it is possible that a more complex set of requirements for placements may improve the effectiveness of the process. While it is not practical to use a student's enrollment in the integrated freshmen engineering sequence or in LWTL to place students (as all engineering students currently enroll in this class) nor ethical to use a student's PES or sex, it may be possible to use student's high school GPA and ACT math score.

For example, new requirements for placing into Pre-Calculus could be more restrictive. Using the equation generated by the linear regression Model 2 and the original placement requirement of needing a 26 or above ACT math score, the new model (Eq. 4.2) suggested that the student with an ACT math score of 26 should also have a high GPA of at least 3.52 in order to make a minimum of a C in the class.

$$\text{Grade of "C" in Pre-Calculus} = -5.17 + 1.41(HSGPA) + .085(26) \quad (4.2)$$

$$2 = -5.17 + 1.41(HSGPA) + .085(26)$$

$$3.52 = HSGPA$$

Another option would be to move the ACT cutoff score to a different number and then also take into account GPA. For example, the national average for a student's ACT math score in 2016 was 20.8 [137]. Rounding up the average, if a student has a score of 21 then it must be accompanied by a GPA of at least 3.82 in order to be eligible to enroll in Pre-Calculus as seen in Eq. 4.3.

$$\text{Grade of "C" in Pre-Calculus} = -5.17 + 1.41(HSGPA) + .085(21) \quad (4.3)$$

$$2 = -5.17 + 1.41(HSGPA) + .085(21)$$

$$3.82 = HSGPA$$

A third option is simply using the equation generated by the analysis. In this case, a student would insert their GPA and ACT math score into the equation, and if the outcome was greater than 2 (the model indicating that the student is predicted to pass Pre-Calculus with a C or higher) then the student could take Pre-Calculus. For instance, a student A with a GPA of 3.8 and an ACT math score of 24 would generate a score of 2.23 and therefore be eligible to take the class. However, student B with a 3.8 GPA and a lower score of 19 on the math portion of the ACT would not be able to enroll in Pre-Calculus. In this case, the student scored a 1.80 (calculations shown in Eq. 4.4 and Eq. 4.5).

Student A

$$\text{Predicted Grade in Pre-Calculus} = -5.17 + 1.41(3.8) + .085(24) \quad (4.4)$$

$$\text{Predicted Grade in Pre-Calculus} = 2.23$$

Student B

$$\text{Predicted Grade in Pre-Calculus} = -5.17 + 1.41(3.8) + .085(19) \quad (4.5)$$

$$\text{Predicted Grade in Pre-Calculus} = 1.80$$

Of course, there are drawbacks to using high school GPA and ACT score to place students. A course taken at one high school may be more rigorous than the same course taken at a different high school, even for honors or AP classes [138]. Therefore, students from different high schools could perform similarly but have different high school GPAs. Furthermore, students from smaller schools do not have all the opportunities as students from larger schools, particularly concerning AP courses.

Another potential disadvantage is the use of ACT scores. In relation to college readiness standards, Asian American and White students are more likely to meet benchmarks while African American students are the least likely [137]. American Indian and Hispanic students are also less likely to meet standards. Therefore, using high school GPA and ACT scores to place students could potentially hinder certain students.

4.7.3 Engineering Problem Solving I Models

The third and fourth research questions asked if the engineering class grades were influenced by *ACT math score* and other variables. These questions were asked because if the factors that influence Pre-Calculus grades are used to place students, then are the requirements to enroll in Pre-Calculus the same ones that should be used to enroll in the engineering class? For this analysis, linear regression with forward selection was again implemented with the outcome variable being *grade in Engineering Problem Solving I* and the regressors being the same as the Pre-Calculus model's variables. (Again, *high school rank* was excluded due to high correlation with *high school GPA* and explaining less variance than *GPA* in the model.)

As seen earlier in Table 10, *ACT math score* was a significant influence on both Pre-Calculus and Engineering Problem Solving I grades.

The first variable added to the model was *high school GPA*. It explained 14.0% of the variance and was extremely significant. Generating the same results as Pre-Calculus, the second variable added was *ACT math score*. The adjusted R^2 increased to 18.0% when this term was added. *Enrollment in LWTL* was the third variable included in the model and again increased the amount of variance explained. The fourth term added was *ACT science score*. After adding this term, however, the adjusted R^2 did not increase more than half of a percentage point. The results of these steps are shown in Table 15.

Table 15 Engineering Models for Achievement Study

Model	Coefficient	Std. Error	Conf. Interval		Adjusted R^2	p-value
			2.5%	97.5%		
Engineering Model 1					14.0%	
HSGPA	1.37	0.072	1.232	1.516		< 2e-16 ***
Engineering Model 2					18.0%	
HSGPA	1.09	0.076	0.947	1.243		< 2e-16 ***
ACT M	0.082	0.008	0.066	0.097		< 2e-16 ***
Engineering Model 3					19.9%	
HSGPA	1.09	0.748	0.942	1.236		< 2e-16 ***
ACT M	0.089	0.008	0.067	0.098		< 2e-16 ***
LWTL	-0.431	0.059	-0.547	-0.314		4.91e-13 ***
Engineering Model 4					20.3%	
HSGPA	1.06	0.075	0.916	1.210		< 2e-16 ***
ACT M	0.064	0.009	0.046	0.083		8.63e-12 ***
LWTL	-0.442	0.059	-0.557	-0.326		1.19e-13 ***
ACT S	0.031	0.009	0.014	0.049		0.00046 ***

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

The best model including variables without interaction was Model 3 as it explained the most variance with the fewest variables. The next step dictated that the interaction between the remaining variables be added. Forward regression was again utilized for this process. The interaction between *high school GPA* and *ACT math score*

was the first to be added. This interaction was significant, but rendered *high school GPA* insignificant (Table 16).

Table 16 Engineering Model with Interaction Term

Model	Coefficient	Std. Error	Conf. Interval		Adjusted R ²	p-value
			2.5%	97.5%		
<i>Engineering Model 5</i>					20.4%	
HSGPA	-0.817	0.486	-1.771	0.134		0.09196 .
ACT M	-0.183	0.067	-0.315	-0.051		0.00654 **
LWTL	-0.441	0.059	-0.557	-0.325		1.14e-13 ***
HSGPA*ACT M	0.075	0.019	0.038	0.112		7.24e-05 ***

. p<0.1, *p<0.05, ** p<0.01, *** p<0.001

Although adding this interaction term to the model increased the amount of variance explained, the addition also changed the signs for two of variables (*high school GPA* and *ACT math score*) and lessened their significance to the model. With these results, it was decided that the best model that explained the most variance with the fewest significant variables was Model 3. So the final results for the variables that best predict engineering grade include *high school GPA*, *ACT math score*, and *LWTL*. Again, backward elimination and stepwise regression were also implemented and all the results indicated the same conclusion.

4.7.4 Discussion of Engineering Course Results

Comparing the two classes, Pre-Calculus and Engineering Problem Solving I were indeed similar as both models included *high school GPA*, *ACT math score*, and *LWTL*; the difference between the two models was that two additional variables were included in the Pre-Calculus model. The final step for this study was to determine if

students who were predicted to pass Pre-Calculus would also be predicted to pass the engineering course.

Using Student A and Student B mentioned earlier (on page 59) a model for predicting grade in engineering was implemented again using *ACT math score* and *high school GPA* (Eq. 4.6). Like the Pre-Calculus model, the variable concerning enrollment in either the integrated freshmen engineering sequence or LWTL was not used to make a prediction for *grade in the engineering class*. Similar to the Pre-Calculus results, Student A is expected to earn a higher grade in the engineering course than Student B as seen in Eq. 4.7 and Eq. 4.8.

$$\text{Equation for Predicted Grade in Engineering} = \quad (4.6)$$

$$-3.72 + 1.09(HSGPA) + .082(ACT M)$$

Student A

$$-3.72 + 1.09(3.8) + .082(24) \quad (4.7)$$

$$\text{Predicted Grade in Engineering} = 2.40$$

Student B

$$-3.72 + 1.09(3.8) + .082(19) \quad (4.8)$$

$$\text{Predicted Grade in Engineering} = 1.98$$

Overall, the results from the model indicated that *ACT math score* is not the only variable that influences freshmen math and engineering grades, and therefore the current placement process should be reviewed. Specifically, the linear regression models suggested that high GPA should also be taken into account in addition to a student's ACT math score when placing freshmen engineering majors into an initial math class.

CHAPTER 5

STUDY TWO: PLACEMENT

5.1 Approach and Justification

For Louisiana Tech University engineering majors, a student's ACT math score is the primary tool used for placement in a first mathematics course. Is the placement process well-designed; do students just below and just above the cutoff do similarly well in Pre-Calculus? Regression discontinuity was utilized in order to determine the impact of placement decisions for engineering students who were and were not required to take a "remedial" math course. Again, for the purpose of this study, remedial math class refers to a course below Pre-Calculus such as trigonometry or college algebra. Students with ACT math scores of 26 or above are automatically eligible to enroll in Pre-Calculus and are considered to be "on track" for graduating.

Ideally, participants are chosen at random for an experiment; in education this is often impossible as the data is usually observational. According to the American Educational Research Association, regression discontinuity is one of four methods suggested if estimating causal effects when the data is observational [139]. The Institute of Education Statistics determined that regression discontinuity was a quasi-experimental method that met the prerequisites of a causal relationship [115]. Other research has also concluded that this method is a strong alternative when completely randomized experiments are not possible and that regression discontinuity results are comparable to

randomized experiment results [119, 121, 140]. Particular to this study, regression discontinuity is an optimal choice because this specific type of analysis takes advantage of the nature of the data – that there are two groups split by a cutoff point (a student's ACT math score in this case).

Additionally, multiple studies have used this type of regression to analyze remediation and placement. Studying the effectiveness of an English remedial program for first-time community college students, Moss and Yeaton used regression discontinuity to discover that participation in the remedial program led to similar academic outcomes for both remedial and non-remedial students [123]. Other researchers used regression discontinuity to determine the impact of placement decisions in preparatory math sequences at nine different community colleges in California using student data from over 150,000 participants between 2001 and 2009 [124]. Their overall results indicated that students initially placed in lower-level math classes had worse educational outcomes when compared to students placed in higher-level classes. However, some of the colleges had small positive or negative impacts of placement decisions which led the researchers to believe these colleges had well-calibrated placement policies. A separate set of researchers also used regression discontinuity to study two-year and four-year college students from Texas and determine the effect of college remediation on academic outcomes [141]. The researchers concluded that there was no difference between students just above and below the cutoff in terms of earning a college degree. Testing the impact of math remediation on a subsequent algebra course, among other academic outcomes, Calcagno and Long too used regression discontinuity

[116]. Therefore, due to the type of data available and the literature suggesting it as a viable and reasonable method, regression discontinuity was used for the second study.

5.2 Predictor and Outcome Variables

Two input variables were required for the placement study. The first was assignment to remedial classes (D); denoted by a 1 if the student is assigned to take a remedial class and 0 otherwise. Assignment is based upon the cutoff rule Louisiana Tech University uses in placing students – a student with an ACT math score of 26 or above is not required to take a lower level math class. The next predictor was a student's transformed ACT math score, r . This variable was centered around the cutoff point, as discussed on page 34. Two other variables were used for the analysis and each will be discussed in more detail in the outcome variable section. Two different equations were used for the second study. The first outcome variable was an indicator of remediation. This variable became an input variable for the second equation. The second outcome variable for this study's model was grade in Pre-Calculus.

For the analysis, all students were grouped together. It should be noted, however, that some participant data were removed due to missing variables. Like study one, 2989 students were initially identified as freshman engineering majors enrolling between 2006 and 2014. A total of 471 participants were removed due to missing ACT score, high school GPA or rank, peer economic status, or grade in Pre-Calculus. For this study students just above and below the cutoff used to place students in Pre-Calculus were included; therefore only students with ACT math scores between 24 and 28 inclusive were analyzed. After removing students with ACT scores out of range, the total number of students was 1147.

5.3 Participants

Similar to study one, the participants for the second study were first-time-in-college (FTIC) freshmen engineering majors, identified by enrollment in UNIV 100, enrolled at Louisiana Tech University from 2006 to 2014. Again, this specific time span was chosen for three reasons – the change in requirements for some freshmen classes that occurred in 2006, most recent data available, and the time period for which PES is available. To identify FTIC freshmen, only students enrolled in a required university seminar class (UNIV 100) between the given years are included in the study. Transfer students are not required to take this class and are therefore not included in the study. Possible majors of the students were: basic, biomedical, civil, chemical, cyber, electrical, industrial, mechanical, and nanosystems engineering.

Two groups of participants were included in the study – engineering students that enrolled in a mathematics class lower than Pre-Calculus and engineering students that directly enrolled in Pre-Calculus. A reminder – for the purpose of this study, a lower level mathematics course (below Pre-Calculus) will be considered “remedial.” Also, using regression discontinuity (RD) to analyze placement required only students right above and below the cutoff placement score to be included in the study; therefore only students with ACT math scores between 24 and 28 were included. For the group required to take remedial courses, the sample includes 497 students and for the on-track group the sample size was 650.

About eighty-three percent of the remedial group identified themselves as White and the remaining students identified themselves as Black/African American, Hispanic, Pacific Islander, or Asian American. Race was not used as a predictor variable in the study, therefore students from all races were included. As the majority of both groups

identified as White, the other options for race will be referred to as Black/Minority. In study one, race was used as a variable to predict grade. As race is only used in study two to determine if the two groups in the study are similar, all students reporting a race were included for the second study.

In the non-remedial group, almost ninety percent of the students identified as White. For both groups, the female population was much smaller than the male population. These descriptive statistics are listed in Table 17.

Table 17 Descriptive Statistics for Study Two Participants

Group	N	White	Black/ Minority	Female	Male
Remedial	497	83.3%	16.7%	16.3%	83.7%
Non-Remedial	650	89.8%	10.2%	17.4%	82.6%

5.4 Method

At Louisiana Tech University an ACT math score of 26 is used to split students into remedial or non-remedial math classes. Four specific requirements must be met in order to use RD. First, a discontinuous jump must be present in the data [117]. For this study, students' ACT math scores provide the discontinuous jump; if their score is between 22 and 25, the student is usually placed in college algebra (Math 101). However, if a student scores a 26 or above, they are given credit for college algebra and placed in Pre-Calculus (Math 240). Students who are “on track” take Math 240 in their first quarter.

The second requirement for RD is that the assignment variable is not caused by the treatment [118]. This means that the assignment variable, in this case *ACT math score*, cannot be caused by the treatment. This requirement is met as the score is

evaluated by a qualified outside source before students begin taking classes at the university.

Another requirement for the method used in study two is that the participants must not be able to manipulate treatment [117]. Particular to this study, this means that students must not be able to choose their ACT score or choose which math class they take regardless of their ACT math score. In meeting a part of the third benchmark, ACT tests are objectively graded and cannot be changed by university personnel. Two other situations can occur as well, some students achieved an ACT score allowing them to enroll in Pre-Calculus but chose not to take the class (called “no-shows”) and other students did not score high enough but were allowed to enroll in Pre-Calculus (called “crossovers”). For Louisiana Tech students, this may occur if a student takes the credit exam and passes or if a student decides a remedial class in their best interest. These cases do not violate the manipulation of treatment- students in the remedial classes are taught from the same curricula no matter the ACT score. Also, the majority of students are placed strictly according to their ACT math score. However, as more than five percent of students in the sample were no-shows or crossovers, *fuzzy* regression discontinuity instead of sharp regression discontinuity was used.

In order to provide valid causal evidence, the fourth requirement is that all students just above and just below the jump must share similar relevant characteristics [117]. To comply with only choosing students above and below the cutoff mark, students with an ACT math score two points above and two points below were chosen. Levin and Calcagno state that only students who share similar academic preparedness and backgrounds should be compared [120]. Therefore, the following were chosen as

relevant characteristics: *high school GPA* and *ACT component scores* (indicators of academic preparedness) along with *sex*, *race*, and *PES* (background characteristics). Recalling information presented in an earlier section, the descriptive statistics for the two groups are presented again in Table 18 along with academic statistics for the groups in Table 19. For the descriptive statistics, both groups are largely White males with a slightly higher percentage of White students for the non-remedial group and a very similar distribution of male to female for both groups. Notice that the ACT component scores are slightly higher for the non-remedial group; however, this is to be expected as students' ACT math score is used to separate the two groups. Although the average PES and high school GPA numbers are slightly higher for the non-remedial group, overall the academic statistics are similar as seen in Table 18 and 19.

Table 18 Descriptive Statistics for Placement Study Participants

Group	N	White	Black/ Minority	Female	Male
Remedial	497	83.3%	16.7%	16.3%	83.7%
Non-Remedial	650	89.8%	10.2%	17.4%	82.6%

Table 19 Academic Statistics for Placement Study Participants

Group	Mean ACT E	Mean ACT S	Mean ACT R	Mean ACT M	Mean PES	Mean HSGPA
Remedial	24.6	24.8	24.7	25.3	66.7%	3.5
Non-Remedial	26.0	26.2	26.3	26.4	68.1%	3.6

5.5 Expected Outcomes

At Louisiana Tech University, a student's ACT math score is the primary tool used to place students in Pre-Calculus. While some research has indicated that

standardized test scores are significant influences to freshmen grades, other studies have concluded that high school academics such as rank and GPA are the best predictors of first-year grades. Therefore, it was expected that the results for the regression discontinuity would show that the cutoff score to enroll in Pre-Calculus is currently too high as *ACT math score* may not be as influential to the final grade in the class as possibly thought when the cutoff was established. However, a substantial percentage of the data includes crossovers and no-shows, which may indicate that other decisions beyond a student's ACT math score have been implemented to help improve student success. Therefore a second option was that the model would indicate that the placement policy is well-designed.

5.6 Limitations

Like study one, the data used in study two came from a single university instead of multiple institutions. Though including data from other universities could have increased the generalizability of the study, this study focused on evaluating the placement policy at a single institution.

As previously mentioned, no transfer students were included in the study as these students are not required to take UNIV 100. Also, for studies one and two another group of students were not included – those who bypassed Pre-Calculus. See the limitations section from study one for more details.

The method used for study two also had limitation as fuzzy instead of sharp regression discontinuity was implemented [142, 118]. Using two sets of equations instead of one, the precision of a fuzzy model will be less than a sharp model. However, utilizing sharp RD when the data is fuzzy would produce bias in the results.

5.7 Results for Study Two: Placement

The research question for the second study of this dissertation was, “*Is the cutoff score, an ACT math score of 26, used to place students in Pre-Calculus at Louisiana Tech University one such that students right above and below the cutoff do similarly well in Pre-Calculus?*” In order to analyze the placement process at Louisiana Tech, fuzzy regression discontinuity was implemented. Before the regression discontinuity model was completed, a linear regression model for each group (differentiated by ACT score) was implemented to determine if a significant discontinuity was possible. For these models, the first group included all students who had ACT scores of 24 or 25. The second group included students with ACT math scores of 26 to 28. For each group, the regressor was *ACT math score* and the outcome variable was grade in Pre-Calculus. Jacob and Zhu suggest completing these regressions before further analysis is done to check if a significant discontinuity is possible [142].

A graphical representation of the two regression models is shown in Figure 1. It displays regression lines for predicting *grade in Pre-Calculus* by *ACT math score*.

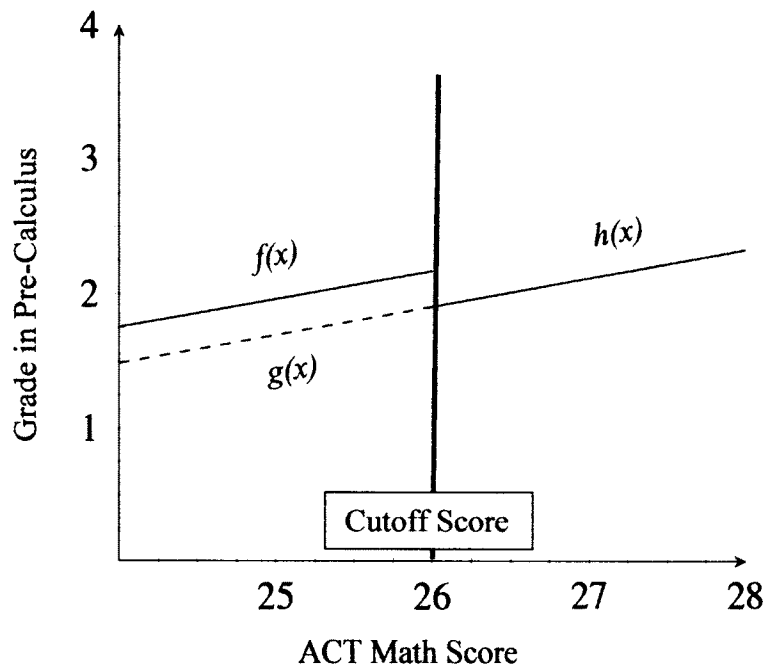


Figure 1: Regression lines of groups by grade in Pre-Calculus and ACT math score (cutoff score = 26). For the figure, $f(x)$ = students with ACT math scores of 24 and 25, $h(x)$ = students with ACT math scores between 26 and 28, and $g(x)$ = projected grades for the second group. Also, recall that grade in Pre-Calculus was transformed into numbers; therefore, 4 represents an “A,” 3 a “B,” and so on.

As *ACT math score* approached 26, the value for predicted grade for the first group approached 2.04. For non-remedial students, the predicted value approached 1.90. The difference between the two values of predicted grade as *ACT math score* approached 26 was 0.14. The gap between the two regression lines indicated that there was a discontinuity. Regression discontinuity was then implemented to determine if the discontinuity was significant and therefore if the ACT cutoff score used to place students in the class was well-designed or not. In that vein, the first equation tested in the study (as seen in Eq. 5.1 and Eq. 5.2) was:

$$T = \alpha_0 + \alpha_1 D + \alpha_1 r + f(r) + \varepsilon_i, \quad (5.1)$$

where initially:

$$f(r) = r * D + r^2 + r^2 * D + r^3 + r^3 * D. \quad (5.2)$$

Before inserting this equation into the next, recall that it must be determined if the interaction and higher order terms are necessary to the model. To evaluate this, regression models were implemented to determine if any of the added terms were significant to the model. The terms r^3 and $r^3 * D$ were not significant and therefore the first to be deleted. Of the remaining terms, $r^2 * D$ was the least significant and therefore the next variable to be removed. After this removal, r^2 was still not significant nor was the interaction between r and D and therefore the final model only included r and D . The final three models of this process are showcased in Table 20.

Table 20 First Stage Regression Discontinuity: Final Three Models

Model	Coefficient	Std. Error	Conf. Interval		p-value
			2.5%	97.5%	
Model 1					
<i>D</i>	0.2302	0.0967	0.0404	0.4199	0.0175 *
<i>r</i>	-0.0967	0.0744	-0.2427	0.0493	0.1942
<i>r * D</i>	0.0452	0.1884	-0.3244	0.4148	0.8105
<i>r</i> ²	0.0144	0.0369	-0.0581	0.0868	0.6975
Model 2					
<i>D</i>	0.2122	0.0612	0.0921	0.3323	0.0005 ***
<i>r</i>	-0.0796	0.0209	-0.1207	-0.0385	0.0002 ***
<i>r</i> ²	0.0058	0.0094	-0.0126	0.0241	0.5361
Model 3					
<i>D</i>	0.2320	0.0521	0.1297	0.3343	9.41e-06 ***
<i>r</i>	-0.0739	0.0188	-0.1107	-0.0370	8.94e-05 ***

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

From the results of these steps, the first stage equation for the model (Eq. 5.3) was:

$$T = 0.33962 - 0.07385r + 0.23203D. \quad (5.3)$$

The second step for regression discontinuity is to substitute Eq. 5.3 into Eq. 5.4:

$$Y = \beta_0 + \beta_1 T + \beta_1 r + f_2(r) + \varepsilon_2, \quad (5.4)$$

where initially $f_2(r)$ was defined as seen in Eq. 5.5:

$$f_2(r) = r * T + r^2 + r^2 * T + r^3 + r^3 * T. \quad (5.5)$$

As with the first stage equation, the second stage of the equations must also be checked to determine if the higher order and interactions terms significantly add to the model. If these terms are not significant, the variables should be removed. Again, higher order and interaction terms were not significant to the model and were the first to be removed. The final equation for the model (Eq. 5.6) was:

$$Y = 1.6969 + 0.6177T + 0.2546r. \quad (5.6)$$

The final three models from the second stage of the regression can be seen in Table 21.

Table 21 Second Stage Regression Discontinuity: Final Three Models

Model	Coefficient	Std. Error	Conf. Interval		p-value
			2.5%	97.5%	
Model 4					
<i>T</i>	0.1177	1.2230	-2.2818	2.5172	0.923
<i>r</i>	0.7981	0.9557	-1.0771	2.6732	0.404
<i>r</i> ²	-0.1618	0.2821	-0.7153	0.3917	0.566
<i>r</i> * <i>T</i>	-1.3515	2.3823	-6.0257	3.3226	0.571
Model 5					
<i>T</i>	0.6547	0.7742	-0.8642	2.1736	0.3979
<i>r</i>	0.2598	0.1151	0.0341	0.4856	0.0241 *
<i>r</i> ²	-0.0025	0.0275	-0.0564	0.0514	0.9272
Model 6					
<i>T</i>	0.6177	0.6593	-0.6760	1.9113	0.3490
<i>r</i>	0.2546	0.0999	0.0586	0.4506	0.0109 *

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

The results indicate that the term *T*, the remediation variable, was not significant to the final model (Model 6). The size of the intercept difference between the remedial

and non-remedial groups was not large enough to be significant and therefore no significant discontinuity was found. However, the second term in the final model, r , was a positive and significant influence. As r is the placement variable, derived from a student's ACT math score as discussed on page 34, a coefficient for this term that is close to zero or modestly positive indicates that the placement policy is accurate [125]. It should also be noted that after removing higher order and interaction terms, the final model was a significant regression model with $F=16.69$, $df=2$, and $p<0.001$. Therefore, the requirements for choosing a final model were met: any interaction or higher order terms not significant were removed from the model and the final model itself was significant.

Recall that the research question for this study was, *“Is the cutoff score, an ACT math score of 26, used to place students in Pre-Calculus at Louisiana Tech University one such that students right above and below the cutoff do similarly well in Pre-Calculus?”* The results of the fuzzy regression discontinuity analysis indicated that the current placement policy at Louisiana Tech for entry to Pre-Calculus is adequately placing students.

CHAPTER 6

STUDY THREE: GRADUATION

6.1 Approach and Justification

A logistic regression was used to analyze the relationship between input variables (environmental and background ones) and graduation for students at Louisiana Tech University. This method was chosen based on its use in other similar studies and its appropriateness. Using a logistic regression was appropriate because of the nature of the data tested. *Graduation* can be coded as a dichotomous variable with graduation indicated with a 1 and not graduating indicated with 0. The possible significantly influential variables are the independent ones. As logistic regression requires these two types of variables, this method was suitable.

Similar studies utilized the same method. For example, a study examining over 400 freshmen students in engineering and other majors used logistic regression to determine which variables were predictive of college persistence [143]. Moller-Wong and Eide performed a longitudinal study testing the relationship between graduation in engineering and demographic/background variables and also used logistic regression to analyze the data [12]. Gayles assessed the relationship between first-year GPA and three different variables (*honors graduation*, *six-year graduation*, and *cumulative GPA* at time of graduation) by using logistic, as well as linear, regression [144]. Similarly, French, Immekus, and Oakes used both linear and logistic regression, but studied cognitive and

non-cognitive variables that impacted engineering student's success and persistence [11]. Another group of researchers implemented a logistic regression to determine what aptitude and affective factors were predictive of retention for freshmen engineering students [40]. A study concerning engineering students from nine different universities tested variables that were influential to graduation also used logistic regression [13]. Considering the type of data used in the study along with the literature supporting the use of the method, logistic regression was chosen.

6.2 Predictor and Outcome Variables

The input variables for study three included *high school GPA*, *high school rank*, *state residency*, *math*, *science*, *English*, and *reading ACT scores*, *race*, *sex*, *enrollment in Living with the Lab*, and *peer economic status*. As in the first study, all of the variables were assigned numeric values before the data were analyzed. However, data pertaining to six different groups of students were analyzed for this study: psychology, social science, health, business, education, and engineering. A more detailed explanation of these groups and why they were chosen is given in section 6.3.

Initially, 400 students were included in the psychology group. Some of the sample was deleted because of missing values. Due to lack of high school GPA or rank, six students were removed from the study. Seventy-six were deleted because no high school code (needed to determine PES) was available. As sixteen students were missing one or more ACT component scores, these were also deleted as well as five others missing PES. As mentioned previously on page 21, several categories of race had small sample sizes or did not indicate a race and were removed accordingly – seventeen total.

For the social sciences group, 475 students were identified at the start. Of the group, 113 were removed due to missing high school codes and thirteen were deleted as no high school GPA or rank were available. Nineteen pieces of student data were deleted due to missing ACT component scores. Four more were deleted as no PES was available. Eighteen students were removed as no race was identified or had significantly small categories of race.

The health group began with 463 students and 145 were first removed because no high school code was given. Five students were deleted when no high school GPA or rank were available. Due to missing ACT component scores, 26 pieces of student data were deleted. Eight students were deleted when no PES was available. Nine other students were missing race along with seven more students as those categories of race were extremely small.

The group of business majors began with 1006 students; as no high school code was given, 191 students were removed. Ten more students were deleted as high school GPA and rank were missing. Forty students were missing one or more ACT component scores. Eight more students were removed due to lack of PES along with 45 students that did not identify a race or had a significantly small category of race.

For the education majors, 551 students were identified initially. Of the sample, 157 were removed due to missing high school code. Twelve students were deleted as no high school GPA or rank was given. Thirteen were removed due to missing ACT component scores. Seven more were deleted as no PES was found. Some categories of race had small sample sizes or did not indicate a race – twenty were removed for this reason.

The largest group of students were engineering majors, $n = 1921$. Due to missing high school codes, 216 participants were removed. As no high school GPA or rank were available, 42 more students were deleted. One hundred twenty-four students were removed because ACT component scores were missing. No PES score was found for another 33 students and consequently deleted. The last step removed 142 students due to missing race or insufficient sample size.

Similar to the first study, the majority of variables were representative of prior knowledge, but were also proven through other works to be potential predictors of graduation for engineering students. As previously mentioned in the literature review, a number of studies concluded that high school academics were positively associated with graduation [10, 28, 40, 13]. *Standardized test scores* have also been proven to be significant influences on the ultimate success of engineering students [13, 28]. Other work has also proven *sex* to be significant, usually indicating that being female makes one less likely to persist while other studies have found that *sex* was not significant [29, 30, 31, 12]. *Race*, *state residency*, and *economics status* have also been proven to be significant influences on graduation [12, 13, 145, 146, 130].

The outcome variables for the graduation study were either *graduation* (denoted with a “1”) or *failure to graduate* (denoted with a “0”). Graduation had to occur within six years of the student’s start data at the university and the resulting degree had to be in the same type of major in which the student originally began.

6.3 Participants

The participants of the third study consisted of all FTIC freshmen (identified via enrollment in UNIV 100) who started at Louisiana Tech University in the fall of one of

the following years: 2006, 2007, 2008, 2009. This group of students was chosen for two specific reasons. First, a six-year graduation timeline was implemented for the study. As data were available through the 2015 school year, the 2009 cohort was the last group with sufficient data to calculate a six-year graduation rate. Therefore, the students were followed from the fall of their start year until their graduation if the event occurred within six years as suggested by previous literature as the most appropriate graduation rate [13]. Louisiana Tech's graduation rate for freshmen cohorts from 2004 to 2011 was 27 percent for a four-year graduation, but jumped to 49 percent at the six year mark [147]. Secondly, the freshmen requirement for Pre-Calculus was changed starting in 2006. At the university, a student must be enrolled in Pre-Calculus in order to also enroll in the freshmen engineering class (Engineering Problem Solving I). Therefore, it was determined that the study should only include students who started in the fall of 2006 until 2009.

The participants were split into six different groups according to the major declared when they began their studies. The first five major categories are reflective of the majors with the largest percentage of bachelor degrees according to the Condition of Education 2014 [148]: psychology, social sciences and history, business, health professions, education. The sixth group is engineering. Engineering students were not contrasted against other math and science majors because of insufficient sample size. Classification of Instructional Programs (CIP) codes were used to determine which majors offered at the university were included in each category. Not all degree programs at Louisiana Tech fell into one of these categories- notably Mathematics, Physics and

Chemistry. The list of specific majors that fall under each major category is given in Appendix A.

For the final analyses, 3269 FTIC freshmen students were identified that enrolled in the fall quarters between 2006 and 2009, declared majors within the aforementioned categories, and were not missing any predictor variables. A total of 280 students majored in psychology, 308 in social sciences/history, 263 in health-related majors, 712 in business, 342 in education, and 1364 in engineering. High school academic averages as well as average PES and state residency are shown in Table 22.

Table 22 Mean GPA, Rank, and ACT by Major

Major	GPA	Rank	State	PES	ACT			
					<i>S</i>	<i>M</i>	<i>E</i>	<i>R</i>
Psyc.	3.31	65.9%	96.1%	72.7%	22.5	21.0	24.0	24.0
S.S.	3.21	62.7%	91.9%	71.2%	22.4	21.0	23.0	23.9
Health	3.48	74.9%	94.3%	71.2%	22.2	21.5	23.8	23.5
Bus.	3.40	70.5%	92.0%	74.3%	22.4	22.2	23.1	23.2
Ed.	3.37	67.8%	95.9%	76.0%	21.4	20.3	23.2	22.9
Engr.	3.44	72.5%	91.1%	75.2%	24.7	25.1	24.6	24.8

Like the first study, at the time of data collection three variables were confounded into one - race, ethnicity, and international status. The students self-selected from the following options: White, Black/African American, Hispanic, Pacific Islander, Asian American, non-resident alien, and decline to identify. The records indicate that the percentage of students at Louisiana Tech who enrolled between 2006 and 2009 and majored in one of the aforementioned categories for each race were as follows: 76.4 % White, 16.3 % Black/African American, 1.9 % Hispanic, and 1.2 % Asian. Less than one percent of students were Pacific Islander. Almost four percent of the population declined

to report a race or were non-resident alien. Unfortunately, due to small percentages of the race categories mentioned above, the study could only include White or Black/African American students. As for the sex of the students, about forty percent of the students were female and sixty percent were male. A breakdown of percentages, after excluding some categories of race, for each major group is given in Table 23.

Table 23 Study Three Sample by Race and Sex

Major (Abbreviation)	N	White	Black	Female	Male
Psychology (Psyc.)	280	73.7 %	26.3 %	68.7 %	32.3 %
Social Sciences (S.S.)	308	77.3 %	22.7 %	42.4 %	57.6 %
Health (Health)	263	86.4 %	13.6 %	81.8 %	18.2 %
Business (Bus.)	712	79.5 %	20.5 %	42.8 %	57.2 %
Education (Ed.)	342	91.3 %	8.7 %	97.7 %	2.3 %
Engineering (Engr.)	1364	86.9 %	13.1 %	14.4 %	85.6 %

6.4 Method

Logistic regression was implemented for the third study to determine the probability that an engineering student would graduate within six years. The best model was chosen based on significance of variables and the adequacy of the model as represented by the HL statistic. Logistic regression was again chosen to represent probability of *graduation* for the following other types of majors as well: psychology, social science, health, business, and education.

6.5 Expected Outcomes

Drawing from previous literature it was expected that all majors would show a positive relationship between *graduation* and *high school GPA* [149]. Extrapolating from studies by Moller-Wong and Eide as well as Thompson and Bolin, it is also expected that the engineering student model would indicate a strong relationship between *graduation*

and *high school rank* [10, 12]. Again, only for engineering disciplines, the researcher suspected that *ACT math* and *science scores* would also show a strong, positive relationship with *graduation*. At many engineering programs, including the ones at Louisiana Tech University, the students must have a certain ACT math score to enroll in certain math and engineering classes or be required to take remedial math classes. Other majors, such as education and psychology, do not require such standards. This led to a hypothesis that *ACT math score* would be influential for engineering students, but not other students. As for students' ACT science scores, this does not directly affect enrolling in classes, but it seemed likely that engineering students would also show an affinity for the subject.

6.6 Limitations

The limitations of study three are similar to the limitations of study one and therefore a more detailed explanation of limitations can be seen on page 46 and 47. The data analyzed all came from a single university instead of multiple institutions. Due to small sample sizes, only two races were included in studies one and three – White and Black. The data used for studies one and three did not contain variables such as *marital status*, *self-efficacy*, and *transfer credit/dual enrollment*, which may influence first-year grades of freshmen students and even graduating with an engineering degree [12, 39, 130, 131]. Again, measures of these factors were unavailable to the researcher. Also, no transfer students were included in the study.

Innate shortcomings of *PES* were also a limitation of the study as well as the type of method used. Observations that contained a missing variable were not able to be processed and therefore were removed from the study, allowing only a subset of student

information to be analyzed. In conjunction with methods, the Hosmer-Lemeshow statistic also has drawbacks. Though widely used, the statistic has limited power and poor interpretability [150, 151].

6.7 Results for Study Three: Graduation

The purpose of this study was to determine factors that influence graduation in engineering and which of those factors are unique to engineering. In that vein, a logistic regression analysis was implemented for each group of students testing the significance of predictor variables against graduation in each specific major category so that significant influential factors for engineering students could be discovered as well as factors for other majors in order to compare and contrast them against engineering. The potential predictor variables were: *high school GPA*, *high school rank*, *state residency*, *sex*, *race*, *PES*, *ACT English score*, *ACT reading score*, *ACT math score*, and *ACT science score*. (It may be noted that the curriculum variable, *LWTL*, is not mentioned in this list; this variable was only regressed against engineering student data and the results concerning the curriculum variable will be discussed after each logistic regression model is completed for each type of major.)

Like study one, before analysis for each group was completed, a correlation matrix was analyzed to ensure that no two variables were highly correlated (reference page 48 for more details concerning correlation).

The matrix showed that moderate correlations existed between different ACT component scores for each group of majors. None of the correlation coefficients explaining the relationship between ACT variables were above 0.8. However, the

correlation coefficients between *high school GPA* and *high school rank* for almost every group of majors exceeded 0.8 as seen in Table 24.

Table 24 Correlation Coefficients between High School GPA and Rank

Major	r
Psyc.	.81
S.S.	.78
Health	.86
Bus.	.83
Ed.	.82
Engr.	.83

To verify that the variables are correlated as well as decide which variable to drop from the study, two additional steps were taken. First, using R studio, an estimated condition number was determined from a logistic regression model with *graduation* as the outcome and input variables of *high school GPA* and rank for each type of major. A condition number, k , is an indicator of the level of correlation in the data. If the value is less than 100, correlation between the variables is not an issue. A number between 100 and 1000 indicates moderate to strong correlation while a condition number above 1000 indicates severe correlation [113]. The estimated condition number, obtained from R Studio, for each group is presented in Table 25. As seen from the table, k is large for all six cases indicating *high school rank* and *high school GPA* are strongly correlated.

Table 25 Condition Number for Each Major Type

Major	k
Psyc.	1026.1
S.S.	866.1
Health	1284.7
Bus.	1108.7
Ed.	1179.0
Engr.	1293.1

The second step involved testing three linear regression models for each type of major where the outcome variable was graduation and the input variables were *high school GPA*, *high school rank*, and both variables respectively. From these models, it was possible to observe if a change in sign occurred from the individual regression models to the combined model where both variables were present. For each type of major, *high school GPA* and *rank* were all positive when individually regressed about graduation. Also of note, *high school GPA* was significant when regressed against graduation individually for every group except social science majors while *high school rank* was significant for four groups, social science and business majors not included. However, when both variables are included in the model the sign of one of the variables changes from positive to negative along with the level of significance. This result again indicated a high correlation such that one of the variables should be removed. The outcomes of these models can be seen in Table 26.

Table 26 High School GPA and Rank Models

Model	Coeff.	p-value
<i>Psychology</i>		
HSGPA	1.1835	0.0187 *
HSRANK	-0.0055	0.5340
<i>Soc.Sci.</i>		
HSGPA	-0.0328	0.938
HSRANK	0.0061	0.429
<i>Health</i>		
HSGPA	-0.5115	0.5457
HSRANK	0.0372	0.0532
<i>Business</i>		
HSGPA	0.8335	0.0151 *
HSRANK	-0.0081	0.2011

Table 26 (Continued)

Model	Coeff.	p-value
Education		
HSGPA	1.9446	0.0002 ***
HSRANK	-0.0114	0.2174
Engineering		
HSGPA	2.2089	2.74e-13 ***
HSRANK	-0.0010	0.8530

p<0.1, *p<0.05, ** p<0.01, *** p<0.001

In determining which variable to keep, it was noted that for four of the six models, *high school GPA* was more significant than *high school rank*. In all four of these cases, *high school GPA* was also a significant variable to the model. Also, when individually regressed *high school GPA* was significant for more of the models than *high school rank*. Given this information, it was decided that *high school rank* should be removed from future regressions.

The first set of logistic regression models involved engineering majors at Louisiana Tech University. The first method implemented was forward selection. The first variable chosen was *high school GPA* followed by *ACT math score* and then *ACT reading score*. All three of the variables were significant to the model and produced odds ratio confidence intervals that did not include 1 (though it should be noted that some odds ratios were very close to 1). Also of note, the Hosmer-Lemeshow (HL) statistic improved with each variable addition. The fourth variable added to the model was *PES*; again, all variables were significant and the model produced a smaller HL statistic. The fifth variable suggested to be added to the model was *sex*. As seen below, adding this variable increased the HL statistic instead of decreasing it. Also, while all other variables remained significant the new variable was not significant. With these results, it was

decided that Model 4 was the best option. To verify this choice, backward elimination and stepwise regression were also implemented. These methods produced the same results as forward selection. The details of each model discussed above are presented in Table 27.

Table 27 Engineering Models

Model	Coeff.	2.5%	97.5%	Odds Ratio	Conf. Interval		p-value	HL
					2.5%	97.5%		
								19.80
								p=0.01
Model 1								
HSGPA	2.1641	1.8190	2.5226	8.7065	6.1658	12.461	<2e-16 ***	6.32
								p=0.61
Model 2								
HSGPA	1.8413	1.4682	2.2272	6.3048	4.3414	9.2740	<2e-16 ***	5.65
ACT M	0.0673	0.0348	0.1002	1.0696	1.0354	1.1054	5.34e-05 ***	
								p=0.69
Model 3								
HSGPA	1.9609	1.5766	2.3590	7.1060	4.8387	10.580	<2e-16 ***	3.46
ACT M	0.0934	0.0571	0.1303	1.0979	1.0588	1.1392	5.55e-07 ***	
ACT R	-0.0496	-0.0794	-0.0200	0.9517	0.9237	0.9802	0.0011 **	
								p=0.90
Model 4								
HSGPA	2.0317	1.6413	2.4365	7.6271	5.1620	11.433	<2e-16 ***	4.97
ACT M	0.0837	0.0466	0.1213	1.0873	1.0477	1.1290	1.1e-05 ***	
ACT R	-0.0509	-0.0809	-0.0213	0.9504	0.9223	0.9789	0.0008 ***	
PES	0.0104	0.0027	0.0182	1.0105	1.0028	1.0184	0.0082 **	
								p=0.76
Model 5								
HSGPA	2.1000	1.7012	2.5137	8.1659	5.4806	12.350	<2e-16 ***	4.97
ACT M	0.0795	0.0421	0.1173	1.0827	1.0430	1.1245	3.45e-05 ***	
ACT R	-0.0482	-0.0783	-0.0184	0.9529	0.9247	0.9818	0.0016 **	
PES	0.0109	0.0032	0.0188	1.0110	1.0032	1.0189	0.0058 **	
SEX	-0.3331	-0.6971	0.0221	0.7167	0.4980	1.0223	0.0691 .	

. p<0.1, *p<0.05, ** p<0.01, *** p<0.001

As expected for all categories of majors, *high school GPA* was a positive and significant influence on graduation for engineering majors. The second variable added to

the model was *ACT math score* which was also a positive and highly significant factor. Again this outcome was anticipated as literature has demonstrated that math preparation is vital to the success of engineering students. A measure of economic status, *PES* was also a positive and significant variable for the model – yet another unsurprising find as other studies have indicated a connection between economic status and graduation. On the other hand, the fourth variable that is a part of the model was a surprise – *ACT reading score*. A significant but negative influence, *ACT reading score* was unexpected as this particular component score is not used in placing students in classes and furthermore the negative coefficient indicates that a higher score hinders a student's chance of graduating. This, seemingly, is contradictory to a logical assumption that if *ACT reading score* is significant for engineering majors then the students with higher reading scores are more likely to understand word problems and write technical reports required in engineering courses, and therefore be more likely to pass these courses and eventually graduate. However, more information was gained when *ACT reading score* was regressed, by itself, against the outcome. As seen in Table 28, this regression showed that the score should have a positive relationship with graduation. Therefore, it is likely that interaction between variables in Model 4 changed the coefficient for this particular variable. Using all four variables (*high school GPA*, *ACT math score*, *ACT reading score*, and *PES*) to predict the probability of an engineering student graduating, a negative coefficient for *ACT reading* is suggested; this result does not, however, mean that the true relationship between *graduation* and *ACT reading score* is negative.

Table 28 Engineering Model for ACT Reading

	Coeff.	2.5%	97.5%	Odds Ratio	Conf. Interval		p-value	HL
					2.5%	97.5%		
Model 6								4.41
ACT R	0.0436	0.0203	0.0670	1.0445	1.0205	1.0693	0.00025 ***	p=0.82

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

Another surprising result was that *ACT science score* did not appear as a significant predictor in the model. It is possible this was due to the collinearity of *ACT math* and *ACT science*, noted earlier. When entered into the model as a lone predictor, *ACT science score* had a highly significant and positive effect. A lack of relation between *race* and *graduation* was a finding that does not concur with previous research [12]. A possible explanation for this is because *race* and *ACT scores* were confounded. The differences in ACT scores between Black and White students, using the engineering data, can be seen in Table 29. The average score for White students in the sample was more than three points higher than for Black students on each part of the ACT.

Table 29 Average ACT Scores by Race for Engineering Students

Race	M	E	R	S
Black	21.3	21.7	21.9	21.6
White	25.6	25.0	25.2	25.2

The second set of regression models were from the psychology major data. Using forward selection, the first variable chosen for the model was *high school GPA*. The second variable added was *ACT science score*. For both models, *GPA* was significant and *ACT science score* was almost significant for the second model. The third variable to be included was *race*. At this point all three variables in the model were significant and

the HL statistic was lower for Model 3 than Model 2 or 1. The fourth variable to be added was *sex*; this addition increased the HL statistic and only *high school GPA* became the only variable significant at $p < 0.05$. After reviewing these results, Model 3 was chosen as the best model. Backward elimination and stepwise regression were also implemented. These methods indicated the Model 3 was best and verified the researcher's selection. The results of the models are described in detail in Table 30.

Table 30 Psychology Models

Model	Coeff.	2.5%	97.5%	Odds Ratio		Conf. Interval		p-value	HL
						2.5%	97.5%		
Model 1									10.42
HSGPA	0.9328	0.3580	1.5310	2.5417	1.4304	4.6227	0.0018	**	p=0.24
Model 2									12.98
HSGPA	0.7524	0.1415	1.3839	2.1222	1.1520	3.9903	0.0173	*	p=0.11
ACT S	0.0592	-0.0099	0.1301	1.0610	0.9901	1.1389	0.0961	.	
Model 3									9.93
HSGPA	0.7964	0.1798	1.4352	2.2175	1.1969	4.2004	0.0126	*	p=0.27
ACT S	0.0835	0.0100	0.1595	1.0870	1.0100	1.1730	0.0395	*	
RACE	0.6296	0.0308	1.2326	1.8769	1.0313	3.4300	0.0280	*	
Model 4									14.14
HSGPA	0.9179	0.2643	1.5952	2.5040	1.3026	4.9292	0.0067	**	p=0.07
ACT S	0.0690	-0.0085	0.1490	1.0714	0.9921	1.1607	0.0847	.	
RACE	0.5951	-0.0079	1.2017	1.8133	0.9915	3.3256	0.0531	.	
SEX	-0.3317	-0.9136	0.2511	0.7177	0.4011	1.2854	0.2629		

. $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Similar to the engineering model, *high school GPA* was again a positive, significant influence on graduation for psychology students. Unlike engineering, *ACT science score* as well as *race* were also positive and significant variables for the model – two results that were expected for engineering but did not occur.

The third set of regression models were for social science majors. The first variable included was *state residency*. *ACT English score* was the second variable added. Using forward selection, the best model only included *state residency* as *ACT English score* was not significant to the model, but produced an HL statistic indicating the model is a good fit.

Backward elimination and stepwise regression were also implemented, but concluded different results. These methods suggested also adding *ACT science score* and *high school GPA* to the model. As seen in Table 31, *ACT science score* was added first; with this addition, *ACT English score* became significant while state residency and ACT science score both produced p values not significant. Also, the HL statistic dramatically increased in Model 3. Model 4 included *high school GPA* along with the previous variables. For this model, only *ACT English* was significant. However, the HL statistic improved to 5.57.

Table 31 Social Sciences Models

Model	Coeff.	2.5%	97.5%	Odds Ratio	Conf. Interval		p-value	HL
					2.5%	97.5%		
								0
<i>Model 1</i>								p=1.0
STATE	-0.7306	-1.6199	0.0999	0.4816	0.1979	1.1050	0.0918	
								5.30
<i>Model 2</i>								p=0.73
STATE	-0.7226	-1.6142	0.1106	0.4855	0.1991	1.1170	0.0964	
ACT E	-0.0317	-0.0802	0.0160	0.9688	0.9230	1.0161	0.1952	
								12.11
<i>Model 3</i>								p=0.15
STATE	-0.7561	-1.6518	0.0811	0.4695	0.1917	1.0845	0.0834	
ACT E	-0.0625	-0.1238	-0.0030	0.9394	0.8835	0.9970	0.0419 *	
ACT S	0.0652	-0.0097	0.1423	1.0674	0.9904	1.1530	0.0912	
								5.57
<i>Model 4</i>								p=0.70
STATE	-0.7405	-1.6394	0.1003	0.4769	0.1941	1.1055	0.0913	
ACT E	-0.0803	-0.1464	-0.0162	0.9228	0.8638	0.9839	0.0153 *	
ACT S	0.0602	-0.0154	0.1379	1.0620	0.9847	1.1478	0.1223	
HSGPA	0.4574	-0.1332	1.0583	1.5800	0.8753	2.8815	0.1313	
· p<0.1, *p<0.05, ** p<0.01, *** p<0.001								

Given these four options of models, none of the models were chosen to represent the probability of graduation for social science majors. Models 1, 2 and 4 showed reasonable fits according to the HL statistic. However, for models 1 and 2, neither of the variables were significant at $p < 0.5$. For Models 3 and 4, not all of the variables included in the model were significant.

None of the variables tested for social science majors exhibited a significant influence on *graduation* if regressed against the outcome by themselves. Therefore, for this study no comparison can be made between social science students and engineering students.

The next set of regression models involved health-related majors. The first variable included in the model was *ACT English score* which was significant. The next variable added was *sex*. The first variable (*ACT English score*) was significant for both models while *sex* was not quite significant ($p = 0.0559$). Therefore, Model 1 was chosen as the best model. The results for each model mentioned are detailed in Table 32.

Table 32 Health Models

Model	Coeff.	2.5%	97.5%	Odds Ratio	Conf. Interval		p-value	HL
					2.5%	97.5%		
Model 1								4.74
ACT E	0.1128	0.0417	0.1871	1.1194	1.0425	1.2057	0.0023 **	p=0.79
Model 2								4.87
ACT E	0.1104	0.0387	0.1852	1.1167	1.0395	1.2035	0.0030 **	p=0.77
SEX	0.9643	0.0570	2.0739	2.6229	1.0586	7.9555	0.0559 .	

. $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The regression for the business major data set indicated that *high school GPA* be the first variable added to the model. It was significant to the model and produced a significant odds ratio. The second variable added to the model was *ACT English score*. For this model, *GPA* was still significant while *ACT English score* was close but not significant. Also, the HL statistic increased from 13.58 to 18.27. Given that the second variable was not significant and the model produced a large HL statistic, Model 1 was chosen as the best model. The results of these steps are shown in Table 33.

Table 33 Business Models

Model	Coeff.	2.5%	97.5%	Odds	Conf. Interval		p-value	HL
				Ratio	2.5%	97.5%		
								13.58
<i>Model 1</i>								p=0.09
HSGPA	0.4672	0.1006	0.8363	1.5955	1.1058	2.3079	0.0127 *	
								18.27
<i>Model 2</i>								p=0.02
HSGPA	0.6230	0.2167	1.0332	1.8645	1.2419	2.8101	0.0028 **	
ACT E	-0.0341	-0.0724	0.0040	0.9665	0.9301	1.0040	0.0803 .	
· p<0.1, *p<0.05, ** p<0.01, *** p<0.001								

. p<0.1, *p<0.05, ** p<0.01, *** p<0.001

Implementing backward and stepwise regression produced slightly different results – both of these methods indicated that *ACT reading score* instead of *ACT English score* should be added to the model after *high school GPA*. However, *ACT reading* was not significant to the model and the HL statistic greatly increased. Furthermore, the odds ratio for *ACT reading* was very close to 1 (shown in Table 34). As this model was not an improvement from model only including *high school GPA*, Model 1 remained the choice for best model.

Table 34 Business Models (Stepwise)

Model	Coeff.	2.5%	97.5%	Odds Ratio	Conf. Interval		p-value	HL
					2.5%	97.5%		
								27.02
<i>Model 3</i>								p<0.01
HSGPA	0.5903	0.1960	0.9887	1.8046	1.2165	2.6878	0.0035	**
ACT R	-0.0299	-0.0651	0.0050	0.9705	0.9370	1.0051	0.0943	.
· p<0.1, *p<0.05, ** p<0.01, *** p<0.001								

. p<0.1, *p<0.05, ** p<0.01, *** p<0.001

Since *ACT English score* was not significant enough to be included in the final model, *high school GPA* was the only significant variable for business majors. Again,

this was an expected result as most studies have found that high school academics are influential for any type of major. This is a similarity between engineering and business majors as *high school GPA* was influential for both groups.

The last group to be analyzed was education majors. Using forward selection, the first variable added to the model was *high school GPA*. It was highly significant and also indicated a significant odds ratio. The second variable added to the model was *PES*, which was also significant. The HL statistic also decreased from Model 1 to Model 2. The last variable suggested to add to the model was *ACT reading score*. However, the increased HL statistic indicated that the model was not a good fit and *ACT reading score* was not significant to the model. Therefore, the model chosen to represent gradation for education majors was Model 2. Backward elimination and stepwise regression agreed on this model as well. The regression models for education are shown in Table 35.

Table 35 Education Models

Model	Coeff.	2.5%	97.5%	Odds Ratio	Conf. Interval		p-value	HL
					2.5%	97.5%		
Model 1								5.10
HSGPA	1.4222	0.8473	2.0246	4.1463	2.3333	7.5732	2.09e-06 ***	p=0.75
Model 2								4.73
HSGPA	1.5367	0.9499	2.1539	4.6494	2.5854	8.6184	5.34e-07 ***	p=0.79
PES	0.0220	0.0079	0.0366	1.0222	1.0079	1.0373	0.00264 **	
Model 3								6.16
HSGPA	1.6379	0.9994	2.3129	5.1444	2.7166	10.104	9.63e-07 ***	p=0.63
PES	0.0228	0.0085	0.0376	1.0230	1.0086	1.0383	0.00204 **	
ACT R	-0.0225	-0.0786	0.0328	0.9777	0.9244	1.0333	0.42669	

p<0.1, *p<0.05, ** p<0.01, *** p<0.001

The education model results showed that these students were most similar to engineering students, sharing two common influential variables – *high school GPA* and *PES*. Though similar, the level of influence of one of the variables differs between the two models. Though the odds ratio for *PES* were very similar for both models, the odds ratio (OR) for *high school GPA* in the education model was 4.65 (if rounding to the hundredth decimal place) while it was 7.63 for the engineering model. Recalling the definition of OR, this means that a one unit increase in *high school GPA* for an engineering student results in the odds of that student graduating increasing by a factor of 7.63. For an education student, a one unit increased in *GPA* improves the odds of graduating by 4.65. In this case, a one unit increased is a point increased in the student's high school GPA. An example is given below for an engineering and education student. First, an equation computing the odds of graduation is given (Eq. 6.1 and Eq. 6.10). Recall all variables except *high school GPA* must stay the same in the second equation and that the *GPA* variable must increase by one point. Random values were chosen for all other variables in the model. The equation for likelihood of graduating for an engineering student is:

$$\ln\left(\frac{\pi}{1-\pi}\right) = -9.81018 + 2.0317(\text{HSGPA}) + 0.0837(\text{ACT M}) - 0.0509(\text{ACT R}) + 0.0104(\text{PES}) \quad (6.1)$$

Equations 6.1-6.5 denote likelihood of graduating in engineering given that a student has a high school GPA of 3.0, an ACT math score of 20, an ACT reading score of 20, and PES value of 20.

$$\ln(\text{odds}) = -9.81018 + 2.0317(3.0) + 0.0837(20) - 0.0509(20) + 0.0104(20) \quad (6.2)$$

$$\ln(odds) = -2.85108 \quad (6.3)$$

$$odds = e^{-2.85108} \quad (6.4)$$

$$odds = 0.0578 \quad (6.5)$$

As seen in Eq. 6.5, the odds of graduating are about 5.8%. The next equation will compute the odds of graduating if the engineering student increased high school GPA to a 4.0. If the odds ratio for *high school GPA* for an engineering student is 7.63 then the odds should increase by a factor of 7.63 (as seen in Eqs. 6.6-6.9).

$$\ln(odds) = -9.81018 + 2.0317(4.0) + 0.0837(20) \quad (6.6)$$

$$-0.0509(20) + 0.0104(20)$$

$$\ln(odds) = -0.81938 \quad (6.7)$$

$$odds = e^{-0.81938} \quad (6.8)$$

$$odds = 0.4407 \quad (6.9)$$

After a one point increase in *high school GPA*, the odds of graduating increase to 44.1% which is about 7.63 times the odds when the student had a GPA of 3.0. The same calculations can be made for an education student using the equation:

$$\ln\left(\frac{\pi}{1-\pi}\right) = -7.19054 + 1.5367(HSGPA) + 0.22(PES) \quad (6.10)$$

If an education student has a PES score of 20 and a high school GPA of 3.0, then the odds of graduating are 0.1176 as shown in Eqs. 6.11-6.14.

$$\ln(odds) = -7.19054 + 1.5367(3.0) + 0.022(20) \quad (6.11)$$

$$\ln(odds) = -2.14044 \quad (6.12)$$

$$odds = e^{-2.14044} \quad (6.13)$$

$$odds = 0.1176 \quad (6.14)$$

The odds ratio for a one point increase in *high school GPA* for an education student is 4.65. Therefore, increasing the GPA to 4.0 should increase the odds by a factor of 4.65 (as seen in Eqs. 6.15-6.18).

$$\ln(odds) = -7.19054 + 1.5367(4.0) + 0.022(20) \quad (6.15)$$

$$\ln(odds) = -0.60374 \quad (6.16)$$

$$odds = e^{-0.60374} \quad (6.17)$$

$$odds = 0.5468 \quad (6.18)$$

As seen from the result, Eq. 6.18, the odds do increase by a factor of 4.65 when there is a one point increase in *high school GPA*. Therefore, an increase in *high school GPA* has a greater influence on the odds of graduating for an engineering student than an education student.

The final logistic regression model chosen to represent each group is shown in Table 36. To answer the first research question of study three (factors that influence engineering student graduation), four of the nine possible variables were influential to an engineering student's graduation – *high school GPA*, *ACT math score*, *ACT reading score*, and *PES*. Not listed is *ACT science score* – a measure suspected to be influential for engineering students and instead was significant in the psychology model. Race was also important to the psychology model and not the engineering one. It is suspected that this measure was not a part of the final engineering model due to the small sample sizes in the data, which involved only including two races, and the fact that Black students only represented 13% of the engineering population. Another variable not seen in the engineering model was *ACT English*. Preliminary results from a precursor study showed that *ACT English score* would be a final predictor for the business model, but also for the

engineering model. However, it seems likely that the correlation between *ACT reading* and *English* played a role in switching which variable was significant for engineering students, though it should be noted that the signs of the coefficients for the models were different.

Table 36 Final Graduation Models

Model	Coeff.	2.5%	97.5%	Odds Ratio	Conf. Interval		p-value		HL
					2.5%	97.5%			
Psychology									9.93
HSGPA	0.7964	0.1798	1.4351	2.2175	1.1969	4.2004	0.0126	*	
RACE	0.6296	0.0308	1.2326	1.8769	1.0313	3.4300	0.0395	*	
ACT S	0.0835	0.0100	0.1595	1.0870	1.0100	1.1730	0.0280	*	
Health									4.74
ACT E	0.1128	0.04165	0.1871	1.1194	1.0425	1.2057	0.0023	**	
Business									13.58
HSGPA	0.4672	0.1006	0.8363	1.5955	1.1058	2.3079	0.0127	*	
Education									4.73
HSGPA	1.5367	0.9499	2.1539	4.6494	2.5854	8.6184	5.34e-07	***	
PES	0.0220	0.0079	0.0366	1.0222	1.0079	1.0373	0.0026	**	
Engineering									3.46
HSGPA	2.0317	1.6413	2.4364	7.6271	5.1620	11.433	<2e-16	***	
ACT M	0.0837	0.0466	0.1213	1.0873	1.0478	1.1290	1.10e-05	***	
ACT R	-0.0509	-0.0809	-0.0213	0.9504	0.9223	0.9789	0.0008	***	
PES	0.0104	0.0027	0.0182	1.0105	1.0028	1.0184	0.0082	**	

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

The second question in the graduation study asked which variables were unique to engineering majors. *High school GPA* was not a unique predictor as it was also significant for psychology, business, and education majors in terms of graduation. This was an expected outcome as multiple studies have previously reported *high school GPA* as a significant influence on graduation for engineering as well as other majors.

A second variable not unique to engineering students was *PES*, as it was also significant for education majors. *ACT math* and *reading scores*, however, were unique to engineering when compared to the other majors in the study; *ACT math* was a positive and highly significant variable while *ACT reading* was also significant but a negative factor when included in the model with *GPA* and *PES*. As predicted, *ACT math score* was influential for engineering students while not significant for other majors. *ACT reading score* was unexpected, but was only significant for the engineering model.

The last research question for this study asked if *enrollment in LWTL* was a significant influence on graduation for engineering students. This variable was not included in the previous engineering model as the research questions one and two focused on the differences or similarities between engineering and other majors. Therefore, two new sets of regression models were implemented. The first model regressed the *LWTL* variable against *graduation* for engineering students. This allowed the researcher to determine if the variable was significant to *graduation* if no other variables were included in the model and the direction of the impact (negative or positive). The second set of models used forward regression to determine if *LWTL* was significant when other variables, such as *high school GPA* and *ACT math score*, were included in the model.

In a model only containing enrollment in the freshmen engineering curriculum, the variable was positive and significant. Though study one found that taking the *Living with the Lab* curriculum was a negative influence for freshmen grades, these results indicated that enrollment was a positive factor in terms of *graduation*. The details of the model can be seen in Table 37.

Table 37 Linear Regression for Graduation versus LWTL Variable

Model	Coefficient	Std. Error	Conf. Interval		Odds Ratio	p-value
			2.5%	97.5%		
<i>Model 1</i>						
LWTL	0.3225	0.1221	0.0825	0.5615	1.3806	0.00828 **

· p<0.1, *p<0.05, ** p<0.01, *** p<0.001

The second set of logistic regression models used forward logistic regression to determine if *LWTL* was significant to *graduation* when the model also included *high school GPA*, *ACT component scores*, *state residency*, *sex*, *PES score*, and *race*. Through the analysis, it was determined that *LWTL* did not significantly add to the engineering graduation model. Using forward selection, the *LWTL* variable was the seventh of ten variables added to the model. Compared to Model 4, which contained all significant variables and an acceptable HL statistic, the model which contained the *LWTL* variable was inferior. Backward elimination along with stepwise regression concluded the same results as forward selection. Therefore, the analysis showed that though *enrollment in LWTL* was a negative influence for Pre-Calculus and Engineering Problem Solving I grades and a positive influence for engineering graduation by itself, this variable did not have a significant impact on *graduation* compared to *high school GPA*, *ACT math score*, *PES*, and *ACT reading score*. The details of the regression models are shown in Table 38.

Table 38 Engineering Models Including the LWTL Variable

Model	Coeff.	2.5%	97.5%	Odds Ratio	Conf. Interval		p-value		HL
					2.5%	97.5%			
									19.80
Model 1									p=0.01
HSGPA	2.1641	1.8190	2.5226	8.7065	6.1658	12.461	<2e-16	***	
									6.32
Model 2									p=0.61
HSGPA	1.8413	1.4682	2.2272	6.3048	4.3414	9.2740	<2e-16	***	
ACT M	0.0673	0.0348	0.1002	1.0696	1.0354	1.1054	5.34e-05	***	
									5.65
Model 3									p=0.69
HSGPA	1.9609	1.5766	2.3590	7.1060	4.8387	10.580	<2e-16	***	
ACT M	0.0934	0.0571	0.1303	1.0979	1.0588	1.1392	5.55e-07	***	
ACT R	-0.0496	-0.0794	-0.0200	0.9517	0.9237	0.9802	0.0011	**	
									3.46
Model 4									p=0.90
HSGPA	2.0317	1.6413	2.4365	7.6271	5.1620	11.433	<2e-16	***	
ACT M	0.0837	0.0466	0.1213	1.0873	1.0477	1.1290	1.1e-05	***	
ACT R	-0.0509	-0.0809	-0.0213	0.9504	0.9223	0.9789	0.0008	***	
PES	0.0104	0.0027	0.0182	1.0105	1.0028	1.0184	0.0082	**	
									4.97
Model 5									p=0.76
HSGPA	2.1000	1.7012	2.5137	8.1659	5.4806	12.350	<2e-16	***	
ACT M	0.0795	0.0421	0.1173	1.0827	1.0430	1.1245	3.45e-05	***	
ACT R	-0.0482	-0.0783	-0.0184	0.9529	0.9247	0.9818	0.0016	**	
PES	0.0109	0.0032	0.0188	1.0110	1.0032	1.0189	0.0058	**	
SEX	-0.3331	-0.6971	0.0221	0.7167	0.4980	1.0223	0.0691	.	
									5.28
Model 6									p=0.73
HSGPA	2.1295	1.7282	2.5459	8.4107	5.6305	12.755	<2e-16	***	
ACT M	0.0850	0.0469	0.1237	1.0887	1.0480	1.1317	1.43e-05	***	
ACT R	-0.0473	-0.0775	-0.0174	0.9538	0.9254	0.9828	0.0020	**	
PES	0.0123	0.0044	0.0203	1.0123	1.0044	1.0205	0.0025	**	
SEX	-0.3594	-0.7250	-0.0027	0.6981	0.4843	0.9973	0.0508	.	
RACE	0.3519	-0.1027	0.7945	1.4218	0.9024	2.2134	0.1233	.	
									1.1
Model 7									p=1.00
HSGPA	2.1275	1.7260	2.5440	8.3938	5.6183	12.731	<2e-16	***	
ACT M	0.0827	0.0443	0.1217	1.0863	1.0453	1.1294	2.75e-05	***	
ACT R	-0.0476	-0.0778	-0.0177	0.9535	0.4809	0.9825	0.0019	**	
PES	0.0124	0.0045	0.0204	1.0125	1.0045	1.0207	0.0022	**	
SEX	-0.3661	-0.7321	-0.0089	0.6934	0.4809	0.9912	0.0470	*	
RACE	0.3482	-0.1062	0.7907	1.4165	0.8993	2.2049	0.1272	.	
LWTL	0.1305	-0.1334	0.3927	1.1394	0.8751	1.4809	0.3307	.	

. p<0.1, *p<0.05, ** p<0.01, *** p<0.001

CHAPTER 7

SUMMARY AND CONCLUSION

7.1 Study One: Achievement

The participants of this study included 2,235 first-time-in-college freshman engineering students. Of this sample, 15.9% were female and 84.1% were male while 12.1% were Black and 87.9% were White. Using linear regression with forward selection and verifying with backward and stepwise regression, three specific conclusions were made concerning Louisiana Tech engineering students and first-year courses. First, *ACT math score* is a positive and significant variable in terms of predicting Pre-Calculus and Engineering Problem Solving I grades. For Pre-Calculus, the regression model explained 13.6% of the variance and 10.2% for the engineering model. Secondly, the best model to predict *grade in Pre-Calculus*, given the available input variables, had five regressors: *high school GPA*, *ACT math score*, the curriculum variable (*enrollment in Living with the Lab*), *peer economic status*, and *sex*. Each of these variables were significant to the model; the majority of the variables were positive influences while enrollment in Living the Lab was negative. Third, the model for predicting grade in the engineering class was similar to Pre-Calculus as *high school GPA* and *ACT math score* were also positive and significant variables in the final model for engineering; also, *enrollment in Living with the Lab* was again a negative influence in the model. The final engineering model did not, however, include *sex* or *peer economic status*.

Four main conclusions were made from these results. One, the variables that most influence *grade in Pre-Calculus* are similar to the ones which influence *grade in Engineering Problem Solving I*, particularly in terms of variables that could be used to place students in an initial math course at the university. Two, more research should be done to determine if the placement requirements for enrolling in Pre-Calculus at Louisiana Tech University are appropriate. Currently, the university mainly relies on a specific ACT math score to place students. The results of this study suggest that *high school GPA* is the best single predictor of grade in Pre-Calculus and perhaps should be used in the placement process if it is determined that the current practice is not appropriate. Three, the conclusions from this study are potentially valuable for recruiters and student success specialists at Louisiana Tech University. Recruiters could use this information to focus on students with high ACT math scores and high school GPAs. Student success specialists could pinpoint freshmen students who may struggle with these classes and prepare to intervene before a student fails a class or drops out of engineering.

Continued analysis of Louisiana Tech engineering student data should be undertaken to determine the accuracy of the equation generated by the linear regression model for predicting *grade in Pre-Calculus* as well as *grade in the engineering course*. Lastly, enrollment in *Living with the Lab* was a negative influence on grade in both classes (Pre-Calculus and Engineering Problem Solving I). Further research should be conducted to determine if this variable also has a negative effect on the graduation of engineering students.

7.2 Study Two: Placement

Using regression discontinuity analysis to evaluate the placement policy for enrolling in Pre-Calculus, this study capitalized on the split nature of the data. Two types of students were included in the study, first-time-in-college engineering freshmen who enrolled in a lower level mathematics course than Pre-Calculus and those who immediately enrolled in Pre-Calculus. Typically, an ACT math score of 26 is used to separate the two groups. However, the sample also included crossovers and no-shows – students who placed in Pre-Calculus and did not enroll as well as students who enrolled in Pre-Calculus yet had an ACT math score below 26. From the analysis, it was concluded that the placement policy at Louisiana Tech was appropriate. The university should continue using an ACT math score of 26 as the cutoff point between remediation and enrollment in Pre-Calculus.

7.3 Study Three: Graduation

Using logistic regression, separate analyses were done to determine what factors influence graduation among psychology, social sciences, health, business, education, and engineering students, respectively. A total of 3,269 participants were included in the study with 280 psychology, 308 social sciences, 263 health, 712 business, 342 education, and 1364 engineering majors between 2006 and 2009. The first part of the research answered two specific questions – what variables are most influential for engineering students and are any of those variables unique to engineering. Concurring with previous research and literature, *high school GPA* held a positive and significant relationship with graduation for the majority of types of majors. It was influential for engineering majors as well as psychology, health, and education majors. Another variable significant, but not

unique, to engineering majors was *peer economic status* which was also influential for education majors. *ACT science score* was not a significance influence for engineering students, but was a positive influence for psychology students. Another unexpected result, *ACT reading score* was a negative influence for engineering majors. *ACT math score* was also significant for engineering majors, but was not significant for any other major.

The second part of this study found that while *enrollment in Living with the Lab* was a positive and significant influence on *graduation* for engineering majors when regressed by itself, this variable was not significant when other variables, such as *high school GPA* and *ACT math score*, were added to the model.

The results of this study are useful in multiple ways. The findings verify conclusions from other studies which state that *high school GPA* is usually a positive and significant predictor of graduation no matter a student's major. Also, other studies have shown that math preparation or knowledge is important for engineering students specifically; again, this study concurs as *ACT math score* was found to be influential for engineering students alone. Additionally, these results denote the differences and similarities of variables which influence graduation for different types of majors which could be useful for recruiting as well as retaining students at Louisiana Tech University.

The study also answered a question about *enrollment in Living with the Lab*. Though this was a negative influence on freshmen grades for engineering students, it was a positive, but not a significant, factor in terms of graduation.

7.4 Future Work and Conclusion

Continued analysis of these topics could be undertaken to gain more knowledge on the success of engineering students. First, these studies used demographic and background variables while other studies have shown that variables such as marital status, self-efficacy, and personality can influence engineering students' grades and graduation. Obtaining and including variables such as these could change the outcomes and/or improve the predictive power of the models.

Another topic of future interest is a more detailed study of factors that affect graduation for engineering students. Are the factors which influence engineering students as a whole the same ones that influence students required to take remedial classes? Do other variables such as amount of time spent studying or type of internship experience affect graduation for engineering students?

Though many questions are still unanswered, the studies included in this work have shown that *high school GPA*, *ACT math score*, and *peer economic status* are positive and significant influences for engineering students in terms of freshmen grades as well as graduation. Also, a regression discontinuity analysis showed that the current placement policy for enrolling students in Pre-Calculus or remedial classes is appropriate. The results from the current work will be useful for recruiting and retaining engineering students at Louisiana Tech University. This, in turn, will assist the nation in producing more engineers to fill the void in the workplace. As more engineering majors are retained and graduate, more engineers will enter the work field where there is currently a high demand for these qualified individuals.

APPENDIX A

STUDY THREE: LIST OF MAJORS

Table 39 Majors within Each Group in Study Three

Category	Major
Psychology	Psychology
Social Sciences	Geographic Information Science Political Science Sociology History
Health	Health Informatics and Information Management Medical Technology Nutrition and Dietetics Nursing Pre-Professional Speech-Language Pathology
Business	Business Administration Management-Business Management and Entrepreneurship Sustainable Supply Chain Management Accounting Business Economics Finance Management- Human Resources Management Computer Information Systems Marketing
Education	Elementary Education (1 - 5) Elementary Education and Special Education Mild/Moderate (1 - 5) Middle School Education Math (4 – 8) Middle School Education Science (4 – 8) Agricultural Education (6 – 12) Business Education (6 – 12) English Education (6 – 12) Family/Consumer Science Education (6 – 12) Mathematics Education (6 – 12) Social Studies Education (6 – 12) Biology Education (6 – 12) Chemistry Education (6 – 12) Physics Education (6 – 12) Earth Science Education (6 – 12) Art Education (K – 12) Health and Physical Education (K – 12) Early Childhood Education (PK - 3)

Table 39 (Continued)

Category	Major
Engineering	Biomedical Engineering
	Chemical Engineering
	Civil Engineering
	Electrical Engineering
	Mechanical Engineering
	Industrial Engineering
	Nanosystems Engineering
	Cyber Engineering

APPENDIX B

APPROVALS

B.1 IRB Approval Letter

LOUISIANA TECH
U N I V E R S I T Y

MEMORANDUM**OFFICE OF UNIVERSITY RESEARCH**

TO: Dr. Marisa Orr and Ms. Sara Hahler
FROM: Barbara Talbot, University Research
SUBJECT: HUMAN USE COMMITTEE REVIEW

DATE: January 13, 2014

In order to facilitate your project, an EXPEDITED REVIEW has been done for your proposed study entitled: "Establishing a Database for Engineering Education Research at Louisiana Tech University"

HUC 1157

The proposed study's revised procedures were found to provide reasonable and adequate safeguards against possible risks involving human subjects. The information to be collected may be personal in nature or implication. Therefore, diligent care needs to be taken to protect the privacy of the participants and to assure that the data are kept confidential.

Mr. Robert Vento is Louisiana Tech's expert on FERPA and student records. See his quote below:

I see no FERPA problems with continuance of this study under the revised purpose(s) defined by HUC 1157 so long as the "...studies are conducted in such a manner as will not permit the personal identification of students and their parents by persons other than the representatives <principals> of such organizations and such information will be destroyed when no longer needed for the purpose for which it is conducted".

This approval is for data extracted from the SCT IA+ 2000 Student Information System (Fall 1999 - present), and the ST legacy headers and transcripts data (Fall 1990 - Summer 1999) for students admitted to, enrolled in, and graduated from COES and it's predecessor, COE during those periods. Protection from unauthorized disclosure of personally identifiable (PID) student and/or parent information is paramount; PID disclosure without written consent is expressly prohibited by law.

As stated in my e-mail of 12-04-2012 to Dr. Orr (referenced in HUC 1157):
Research data requirements for FINAID-type information, Student Accounts data, and Faculty data, require separate review and approval by the respective executive authorities.

If you have participants in your study whose first language is not English, be sure that informed consent materials are adequately explained or translated. Since your reviewed project appears to do no damage to the participants, the Human Use Committee grants approval of the involvement of human subjects as outlined.

Projects should be renewed annually. This approval was finalized on January 10, 2014 and this project will need to receive a continuation review by the IRB if the project, including data analysis, continues beyond January 10, 2015. Any discrepancies in procedure or changes that have been made including approved changes should be noted in the review application. Projects involving NIH funds require annual education training to be documented. For more information regarding this, contact the Office of University Research.

You are requested to maintain written records of your procedures, data collected, and subjects involved. These records will need to be available upon request during the

conduct of the study and retained by the university for three years after the conclusion of the study. If changes occur in recruiting of subjects, informed consent process or in your research protocol, or if unanticipated problems should arise it is the Researchers responsibility to notify the Office of Research or IRB in writing. The project should be discontinued until modifications can be reviewed and approved.

If you have any questions, please contact Dr. Mary Livingston at 257-2292 or 257-5066.

NOTE: Information may be personally identifiable.

B.2 IRB Proposal**HUMAN SUBJECTS APPROVAL PACKET****DEPARTMENT HEAD APPROVAL FORM**

TO: Project Directors

FROM: Barbara Talbot, Office of University Research

btalbot@latech.edu

318-257-5075 phone

318-257-5079 fax

<http://research.latech.edu/>

SUBJECT: HUMAN USE COMMITTEE REVIEW

DATE:

Please submit this page signed by your Department Head or Dean when submitting a proposal to the Human Use Committee for expedited approval.

Their signature is stating that they are aware of this proposal and/or survey being conducted, and all aspects of the study comply with the appropriate University Policies and Procedures.

(print or type below)

Department

Mechanical Engineering

Department Head Name

David Hall

Signature

Date

(Actual original signature required)

<p>Do you plan to publish this study?</p> <p><input type="checkbox"/> <u>YES</u> <input type="checkbox"/> NO</p>
<p>Will this study be published by a national organization?</p> <p><input type="checkbox"/> <u>YES</u> <input type="checkbox"/> NO</p>
<p>Are copyrighted materials involved?</p> <p><input type="checkbox"/> YES <input type="checkbox"/> <u>NO</u></p> <p>Do you have written permission to use copyrighted materials?</p> <p><input type="checkbox"/> YES <input type="checkbox"/> <u>NO</u></p>
<p>COMMENTS:</p>

STUDY/PROJECT INFORMATION FOR HUMAN SUBJECTS

COMMITTEE

<p>Describe your study/project in detail for the Human Subjects Committee.</p> <p>Please include the following information.</p>

TITLE: Establishing a database for Engineering Education Research at
Louisiana Tech University

PROJECT DIRECTOR(S): Dr. Marisa Orr, Assistant Professor

Graduate Researcher: Sara Hahler (skh020@latech.edu)

EMAIL: marisao@latech.edu

PHONE: 318 257-3124

DEPARTMENT(S): Mechanical Engineering

PURPOSE OF STUDY/PROJECT:

With calls from the government to increase the number of STEM (Science, Technology, Engineering, and Mathematics) graduates, research on engineering education has become vitally important to our nation. This study will utilize Louisiana Tech University student data to model student performance, pathways, and persistence with an eye towards enhancing the undergraduate engineering education experience. In addition to academic variables and ACT scores, gender, peer socioeconomic status, and race/ethnicity will be considered. The guiding research questions are:

1. Who do we lose from engineering, when do we lose them, and what are the best predictors for each group?
2. What are the different curricular paths to successful graduation in engineering?
Are they different for students of different academic preparation?

Techniques such as logistic regression, regression discontinuity analysis, and structural equation modeling will be used to investigate student outcomes and identify strengths and opportunities for recruiting and retaining a diverse engineering population. These research goals are consistent with Louisiana Tech University's number one priority: recruiting and retaining a model student body and university community. This

single-institution study will serve as a case study with implications for peer institutions, as well as lay the groundwork for future multi-institution studies.

This project will leverage Dr. Orr's previous experiences with institutional data to build a research database of Louisiana Tech's student record data. This will allow Dr. Orr's research group to study a wealth of educational research questions as well as establish baseline data for evaluating curricular, policy, and pedagogical changes. The resulting database will provide a foundation for Dr. Orr's research and PhD student Sara Hahler's dissertation, tentatively titled, "Mathematics ACT Scores, Freshman Engineering and Mathematics Class Grades, and Graduation in STEM: The Connection Between High School Mathematics and Engineering."

This protocol is intended to cover the establishment, maintenance, and use of a database of Louisiana Tech University student records. The database will be used to research questions regarding academic policies and student outcomes.

SUBJECTS:

Degree-seeking Louisiana Tech University undergraduate students enrolled from approximately 1990 to present.

PROCEDURE:

Dr. Orr will work with the registrar, computing center, and the office of institutional research to collect a snapshot of student data at the end of each quarter from approximately 1990 to present. The exact dates will depend on the accessibility of the data. The data will be organized into a relational database with four types of tables:

Common to all tables: term, year, and unique identifier

- **Student table (one record per undergraduate, degree-seeking student):**

ACT score, date of birth, citizenship, country of origin, visa type, ethnicity, pre-college predicted GPA, gender, home county, home zip code, high school classification, high school, high school GPA, high school class rank, high school class size, matriculation major, SAT math and verbal scores, transfer status, transfer hours, and transfer institution.

- **Term table (per enrolled term per undergraduate, degree-seeking student):**

term major, co-op status, cumulative GPA, cumulative hours attempted, cumulative hours earned, enrollment status, exit code, exit term, exit year, student level, major code, on/off campus, academic standing, term GPA, term hours attempted, and term hours earned.

- **Grad table (per degree per undergraduate student):** degree type, major

- **Course table (per course per undergraduate, degree-seeking student):** AP

credit, course prefix, course abbreviation, course number, course contact hours, course credits, course grade points, course grade, and course method (lecture/lab).

Dr. Orr will de-identify the data and securely store a key for re-identification. The graduate researcher will see only de-identified data. A key for re-identification is required for data updates. The key will be password-protected. Only aggregated data will be disclosed in publications and presentations. No individually identifiable information will be disclosed.

INSTRUMENTS AND MEASURES TO INSURE PROTECTION OF CONFIDENTIALITY, ANONYMITY:

Confidentiality will be maintained because student data will be collected and treated in the same manner that is customary in university coursework. No individually identifiable data are released.

RISKS/ALTERNATIVE TREATMENTS:

By using information that is currently being collected on each participant for institutional research, this research poses no risks to the participants beyond the risks normally associated with research conducted in the Office of Institutional Research.

BENEFITS/COMPENSATION:

The potential benefits of the study may influence general and engineering educational practices, as well as institutional and legislative policies.

SAFEGUARDS OF PHYSICAL AND EMOTIONAL WELL-BEING:

This study involves no treatment or physical contact. All information collected from the institution will be held strictly confidential. No individually identifiable data are released.

WRITTEN INFORMED CONSENT FORM:

We are seeking a waiver of consent.

IRB Latitude to Approve a Consent Procedure that Alters or Waives some or all of the Elements of Consent:

Research in general: an IRB may waive or alter the requirement of informed consent under 45 CFR 46.116(d), provided that the IRB finds and documents that all of the following four conditions are met:

- a. the research involves no more than minimal risk to the subjects;
- b. the waiver or alteration will not adversely affect the rights and welfare of the subjects;

- c. the research could not practicably be carried out without the waiver or alteration; and
- d. whenever appropriate, the subjects will be provided with additional pertinent information after participation.

This research meets the above requirements as follows:

- a. By using information that is currently being collected on each participant for institutional records, this research poses no risks to the participants beyond the risks normally associated with research conducted in the Office of Institutional Research.
- b. There are always FERPA concerns with a data set that collects student data. Will collecting this data affect students' "rights and welfare?" The justification for collecting and analyzing this type of data can be found in 20 USC

§1232g(b)(1)(F):

(b) Release of education records; parental consent requirement; exceptions; compliance with judicial orders and subpoenas; audit and evaluation of Federally-supported education programs; recordkeeping.

(1) No funds shall be made available under any applicable program to any educational agency or institution which has a policy or practice of permitting the release of educational records (or personally identifiable information contained therein other than directory information, as defined in paragraph (5) of subsection (a)) of students without the written consent of their parents to any individual, agency, or organization, other than to the following—

(F) organizations conducting studies for, or on behalf of, educational agencies or institutions for the purpose of developing, validating, or

administering predictive tests, administering student aid programs, and improving instruction, if such studies are conducted in such a manner as will not permit the personal identification of students and their parents by persons other than representatives of such organizations and such information will be destroyed when no longer needed for the purpose for which it is conducted;

We are a group of engineering educators who are doing this research to improve how students learn. This research is all about discovering why students succeed or fail as they maneuver through the engineering curriculum.

- c. This research would be impractical without the waiver of consent. Tracking down every student that has attended Louisiana Tech since 1990 would be cost prohibitive.
- d. Our research is available to students through journal articles, conference proceedings, and by email request.

SUPPLEMENTAL DOCUMENTS:

1. Email from registrar

BIBLIOGRAPHY

- [1] D. Langdon, G. McKittrick, D. Beede, B. Khan and M. Doms, "STEM: Good jobs now and for the future," United States Department of Commerce, Retrieved from <http://www.esa.doc.gov/Reports/stem-good-jobs-now-and-future>, 2011.
- [2] J. Morton, "Engineering skills: The threat from China and India?," in *Paper presented at the meeting of European Engineers Forum*, Hanover, England, 2007.
- [3] National Science Foundation:National Center for Science and Engineering Statistics, "Bachelor's degrees awarded, by major field group: 1966–2012," 2015. [Online]. Available: <http://www.nsf.gov/statistics/2015/nsf15326/pdf/tab5.pdf>.
- [4] K. Rask, "Attritioin in STEM fields at a liberal arts college: The importance of grades and pre-college preferences," *Economics of Education Review*, vol. 29, no. 6, pp. 892-900, 2010.
- [5] S. G. Brainard and L. Carlin, "A six-year longitudinal study of undergraduate women in engineering and science," *Journal of Engineering Education*, vol. 87, no. 4, pp. 369-375, 1998.
- [6] G. Huang, N. Taddese, E. Walter and S. Peng, "Entry and persistence of women and minorities in college science and engineering education," National Center for Education Statistics: U.S. Department of Education, Office of Educational Research and Improvement, 2000.
- [7] E. C. Kokkelenberg and E. Sinha, "Who succeeds in STEM studies? An analysis of Binghamton University undergraduate students," *Economics of Education Review*, vol. 29, pp. 935-946, 2010.
- [8] B. Yoder, "Engineering by the numbers," in *American Society for Engineering Education*, Washinton, DC, 2014.
- [9] M. W. Ohland, S. D. Sheppard, G. Lichtenstein, O. Eris, D. Chachra and R. Layton, "Persistence, engagment, and migration in engineering programs," *Journal of Engineering Education*, vol. 97, no. 3, pp. 259-278, 2008.

- [10] R. Thompson and G. Bolin, "Indicators of success in STEM majors: A cohort study," *Journal of College Admission*, vol. 212, pp. 18-24, 2011.
- [11] B. French, J. Immekus and W. Oakes, "Research brief: An examination of indicators of engineering students's success and persistence," *Journal of Engineering Education*, vol. 94, no. 4, pp. 419-25, 2005.
- [12] C. Moller-Wong and A. Eide, "An engineering student retention study," *Journal of Engineering Education*, pp. 7-15, Jan. 1997.
- [13] G. Zhang, T. Anderson, M. Ohland and B. Thorndyke, "Identifying factors influencing engineering student graduation: A longitudinal and cross-institutional study," *Journal of Engineering Education*, pp. 313-320, Oct. 2004.
- [14] W. LeBold and S. Ward, "Engineering retention: National and institutional perspectives," in *American Society for Engineering Education Annual Conference*, 1988.
- [15] D. Budny, W. LeBold and G. Bjedov, "Assessment of the impact of freshman engineering courses," *Journal of Engineering Education*, vol. 87, no. 4, pp. 405-411, 1998.
- [16] Y. Min, G. Zhang, R. Long, T. Anderson and M. Ohland, "Nonparametric survival analysis of the loss rate of undergraduate engineering students," *Journal of Engineering Education*, vol. 100, pp. 349-373, 2011.
- [17] J. Piaget, *The principles of genetic epistemology*, London: Routledge & Kegan Paul, 1985.
- [18] L. Vygotsky, *Mind in society: The development of higher psychological processes*, M. Cole, V. John-Steiner, S. Scribner and E. Souberman, Eds., Cambridge, MA: Harvard University Press, 1978.
- [19] D. Chaipichit, N. Jantharajit and S. Chookhampaeng, "Development of learning and management model based on constructivist learning theory and reasoning strategies for enhancing the critical thinking of secondary students," *Educational Research and Reviews*, vol. 10, no. 16, pp. 2324-2330, 2015.
- [20] T. Wood, "Alternative epistemologies to practice in education: Rethinking what it means to teach and learn," in *Constructivism in Education*, Hillsdale, NJ, Lawrence Erlbaum Associates, 1995, pp. 331-340.
- [21] L. Campbell and B. Campbell, "Beginning with what students know: The role of prior knowledge in learning," in *101 Proven Strategies for Student and Teacher Success*, Thousand Oaks, CA, Corwin Press, 2009.

- [22] R. Marzano, "The importance of background knowledge," in *Building Background Knowledge for Academic Achievement*, Association for Supervision and Curriculum Development, 2004.
- [23] H. Biemans and P. Simons, "CONTACT-2: A computer assisted instructional strategy for promoting conceptual change," *Instructional Science*, vol. 24, pp. 157-176, 1996.
- [24] D. H. Jonassen and B. L. Gabrowski, *Handbook on individual differences, learning, and instruction: Part VII, Prior knowledge*, Hillsdale, NJ: Erlbaum, 1993.
- [25] R. G. McCullough, "The relationship between reader response and prior knowledge on African American students' reading comprehension performance using multicultural literature," *Reading Psychology*, vol. 34, pp. 397-435, 2013.
- [26] S. Portier and L. Wagemans, "The assessment of prior knowledge profiles: A support for independent learning?," *Distance Education*, pp. 65-87, vol. 16(1), 1995.
- [27] R. Thompson and B. Zamboanga, "Prior knowledge and its relevance to student achievement in introduction to psychology," *Teach Psychology*, vol. 3, pp. 96-101, 2003.
- [28] T. Scott, H. Tolson and T. Huang, "Predicting retention of mathematics and science majors," *Journal of College Admission*, vol. 2014, pp. 20-24, 2009.
- [29] R. Felder, G. Felder, M. Mauney, C. Hamrin, Jr. and E. Dietz, "A longitudinal study of engineering student performance and retention. III. Gender differences in student performance and attitudes," *Journal of Engineering Education*, vol. 84, no. 2, pp. 151-174, 1995.
- [30] F. Smyth and J. McArdle, "Ethnic and gender differences in science graduation at selective colleges with implications for admission policy and college choice," *Research in Higher Education*, vol. 45, no. 4, 2004.
- [31] A. Astin and H. Astin, "Undergraduate science education: The impact of different college environments on the educational pipeline in the sciences," Higher Education Research Institute: Graduate School of Education, University of California, Los Angeles, 1992.
- [32] G. Lichtenstein, A. McCormick, S. Sheppard and J. Puma, "Comparing the undergraduate experience of engineers to all other majors: Significant differences are programmatic," *Journal of Engineering Education*, pp. 305-315, 2010.

- [33] R. Lent, H. Sheu, D. Singley, J. Schmidt and C. Gloster, "Longitudinal relations of self-efficacy to outcome expectations, interests, and major choice goals in engineering students," *Journal of Vocational Behavior*, vol. 73, no. 2, pp. 328-35, vol. 73(2), 2008.
- [34] E. Seymour and N. Hewitt, *Talking about leaving: Why undergraduates leave the sciences*, Boulder, CO: Westview Press, 1997.
- [35] A. Astin and L. Oseguera, "Pre-college and institutional influences on degree attainment," in *College student retention: Formula for student success*, Westport, CT, American Council on Education, Praeger, 2005, pp. 245-276.
- [36] M. Parker, "Placement, retention, and success: A longitudinal study of mathematics and retention," *The Journal of General Education*, vol. 54, no. 1, pp. 22-40, 2005.
- [37] D. Beanland, "Challenges and opportunities facing the education of engineers," Address to the Victoria Division of Engineering Australia, SEG Meeting, Melbourne, Mar 2010.
- [38] R. Suresh, "The relationship between barrier courses and persistence in engineering," *Journal of College Student Retention*, vol. 8, no. 2, pp. 215-239, 2006-2007.
- [39] S. Yoon Yoon, P. Imbrie and T. Reed, "First year mathematics course credits and graduation status in engineering," in *6th First Year Engineering Experience Conference*, College Station, TX, 2014.
- [40] L. Moses, C. Hall, K. Wuensch, K. De Urquidi, P. Kauffman, W. Swart, S. Duncan and G. Dixon, "Are math readiness and personality predictive of first-year retention in engineering?," *The Journal of Psychology*, vol. 145, no. 3, pp. 229-245, 2011.
- [41] A. Cabrera, A. Nora and M. Castaneda, "College persistence: Structural modeling of an integrated model of student retention," *Journal of Higher Education*, pp. 123-139, 1993.
- [42] M. Eimers and G. Pike, "Minority and nonminority adjustment to college: Differences or similarities?," *Research in Higher Education*, pp. 77-97, 1997.
- [43] J. Noble and R. Sawyer, "Alternative methods for validating admission and course placement criteria," *AIR Professional File*, pp. 1-9, 1997.
- [44] B. Patterson and K. Mattern, "Validity of the SAT for predicting first-year grades: 2010 SAT validity sample: Statistical report 2013-2," College Board, 2013.

- [45] W. Camara and G. Echternacht, "The SAT[R] I and high school grades: Utility in predicting success in college," The College Board: Office of Research and Development, July 2000.
- [46] C. Veenstra and G. Herrin, "Using the SAT and ACT scores for placement into engineering freshman classes," in *Proceedings of the 2006 American Society for Engineering Education World Conference*, Washington, DC, 2006.
- [47] S. DeBerard, G. Spielmans and D. Julka, "Predictors of academic achievement and retention among college freshmen: A longitudinal study," *College Student Journal*, vol. 38, no. 1, 2004.
- [48] K. McKenzie, K. Gow and R. Schweitzer, "Exploring first-year academic achievement through structural equation modelling," *Higher Education Research and Development*, vol. 23, no. 1, pp. 95-112, 2004.
- [49] N. Honken and P. Ralston, "Freshmeng engineering retention: A holistic look," *Journal of STEM Education*, vol. 14, no. 2, 2013.
- [50] M. Chemers, L. Hu and B. Garcia, "Academic self-efficacy and first year college performance and adjustment," *Journal of Educational Psychology*, vol. 93, no. 1, pp. 55-64, 2001.
- [51] E. Krumrei-Mancuso, F. Newton, D. Wilcox and E. Kim, "Psychosocial factors predicting first-year college student success," *Journal of College Student Development*, vol. 54, no. 3, pp. 247-266, 2013.
- [52] C. P. Veenstra, E. L. Dey and G. Herrin, "Is modeling of freshmen engineering success different from modeling of non-engineering success?," *Journal of Engineering Education*, vol. 97, no. 4, pp. 467-479, 2008.
- [53] Purdue University, "Mechanical Engineering - Program Map," 2016. [Online]. Available: <https://engineering.purdue.edu/ME/Academics/Undergraduate/index.html>.
- [54] Wright State University, "College of Engineering & Computer Science," 2016. [Online]. Available: <https://engineering-computer-science.wright.edu/advising-and-resources/engineering-mathematics-topics-and-materials>.
- [55] S. Moyo, "A study of the possible existence, causes and effects of the mathematical knowledge gap between high school and first year university mathematics programmes and possible remedies for the situation at UNIVEN: A case study," 2013. [Online]. Available: <http://www.assaf.org.za/files/2010/10/Mathematical-gap.pdf>.

- [56] M. Robinson, "Student enrollment in high school AP sciences and calculus: How does it correlate with STEM careers?," *Bulletin of Science, Technology, & Society*, pp. 265-273, 2003.
- [57] K. Aung, R. Underdown and Q. Qian, "Vertical assessment of math competency of freshmen/sophomore engineering students," in *ASEE Annual Conference and Exposition*, Atlanta, GA, July 2013.
- [58] W. Venable, R. McConnell and A. Stiller, "Incorporating mathematics in a freshmen engineering course," *Journal of Engineering and Applied Science*, vol. 2, pp. 552-556, 1995.
- [59] E. Bowen, J. Prior, S. Lloyd and L. Newman-Ford, "Engineering more engineers- bridging the mathematics and careers advice gap," *Engingeer Education: A Journal of Higher Education Academy*, vol. 2, no. 1, pp. 23-32, 2007.
- [60] M. Coupland, A. Gardner and G. Carmody, "Mathematics for engineering education: What students say," in *Proceedings of the 31st Annual Conference of the Mathematics Education Research Group of Australasia*, Brisbane, 2008.
- [61] Institute of Industrial Engineers Solutions, "Math blocks many engineering students (MathSoft Engineering and Education Inc.) (Brief Article)," 2001. [Online]. Available: <https://www.highbeam.com/doc/1G1-74493908.html>.
- [62] M. Kirschenman and B. Brenner, "Education for civil engineering: A profession of practice," *Leadership Management Engineering*, pp. 54-56, 2010.
- [63] A. Zeidmane and N. Sergejeva, "Indirect impact of mathematic in engineering education," *Engineering for Rural Development*, 2013.
- [64] W. James and K. High, "Freshman-level mathematics in engineering: A review of the literature in engineering education," in *American Society for Engineering Education*, 2008.
- [65] M. Ohland, A. Yuhasz and B. Sill, "Identifying and removing a calculus prerequisite as a bottleneck in Clemson's general engineering curriculum," *Journal of Engineering Education*, pp. 253-257, July 2004.
- [66] A. Medhanie, D. Dupuis, B. LeBeau, M. Harwell and T. Post, "The role of ACCUPLACER mathematics placement test on a student's first college mathematics course," *Educational and Psychological Measurement*, 2011.
- [67] K. Norman, A. Medhanie and M. Harwell, "University mathematics placement tests, student mathematics performance, and high school mathematics curriculum," *PRIMUS*, in press.

- [68] B. Parsad, L. Lewis and B. Greene, "Remedial education at degree-granting post-secondary institutions in fall 2000 (NCES 2004-001)," U.S. Department of Education: National Center for Education Statistics, Washington, DC, 2003.
- [69] J. Reisel, M. Jablonski, H. Hosseini and E. Munson, "Assessment of factors impacting success for incoming college engineering students in a summer bridge program," *International Journal of Mathematical Education in Science and Technology*, vol. 43, no. 4, pp. 421-433, 2011.
- [70] D. Krawczyk and E. Toubassi, "A mathematics placement and advising program," in *Assessment Practices in Undergraduate Mathematics*, Washington, DC, Mathematics Association of America, 1999, pp. 181-183.
- [71] S. Britton, D. Daners and M. Stewart, "A self-assessment test for incoming students," *International Journal of Mathematical Education in Science and Technology*, vol. 38, no. 7, pp. 861-868, 2007.
- [72] J. Cederberg, "Administering a placement test: St. Olaf College," in *Assessment Practices in Undergraduate Mathematics*, Washington, DC, Mathematics Association of America, 1999, pp. 178-180.
- [73] S. Callahan, "Mathematics placement at Cottey College," in *Annual Conference of the American Mathematical Association of Two-Year Colleges*, Boston, 1993.
- [74] W. Donovan and E. Wheland, "Placement tools for developmental mathematics and intermediate algebra," *Journal of Developmental Education*, vol. 32, pp. 2-11, 2008.
- [75] E. Jacobson, "Higher placement standards increase course success but reduce program completions," *The Journal of Higher Education*, vol. 55, no. 2, pp. 138-159, 2006.
- [76] M. Lucas and N. McCormick, "Redesigning mathematics curriculum for underprepared college students," *The Journal of Effective Teaching*, vol. 7, no. 2, pp. 36-50, 2007.
- [77] Louisiana Tech University, "About Tech," 2016. [Online].
- [78] J. D. Nelson and B. Schroder, "Establishing an integrated mathematics, engineering, and science curriculum: Lessons learned," in *American Society for Engineering Education*, Albuquerque, NM, 2001.
- [79] J. D. Nelson, S. Napper, E. Elmore, J. P. Carpenter and B. Deese, "An integrated freshmen engineering curriculum," in *Frontiers in Education Conference*, 1998.

- [80] J. Nelson and S. Napper, "Ramping up an integrated engineering curriculum to full implementation," in *Frontiers in Education Conference*, San Juan, Puerto Rico, 1999.
- [81] D. Hall and M. Barker, "Living with the Lab - Boosting experiential learning and creativity in 1st year engineering students," *Intelligent Automation and Soft Computing*, vol. 13, no. 1, pp. 3-18, 2007.
- [82] D. Hall, J. M. Barker and J. Nelson, "Living with the Lab: Expanding a project-based freshman curriculum to over 350 students," in *American Society for Engineering Education Annual Conference*, Pittsburg, PA, 2008.
- [83] D. Hall, S. R. Cronk, P. D. Brackin, J. M. Barker and K. B. Crittenden, "Living with the Lab: A curriculum to prepare freshman students to meet the attributes of "The Engineer of 2020"," in *American Society for Engineering Education Annual Conference*, Pittsburg, PA, 2008.
- [84] R. A. Long, "Mellon foundation NCES-CEEB crosswalk," 2011. [Online]. Available: <https://engineering.purdue.edu/MIDFIELD/data-archive.htm> .
- [85] G. Chen, "Public School vs. Private School," 3 March 2015. [Online]. Available: <http://www.publicschoolreview.com/blog/public-school-vs-private-school>.
- [86] National Center for Education Statistics, "Digest of Education Statistics, 2011 (NCES 2012-001, Table 3)," National Center for Education Statistics, Washington, DC, 2012.
- [87] United States Department of Agriculture, "National School Lunch Program," 1 Jan. 2016. [Online]. Available: <http://www.fns.usda.gov/nslp/national-school-lunch-program-nslp>.
- [88] National Center for Educational Statistics, U.S. Department of Education, "Characteristics of public and private elementary and secondary schools in the United States: Results from the 2011-12 schools and staffing survey," 2013. [Online]. Available: <http://nces.ed.gov/pubs2013/2013312.pdf>.
- [89] M. Harwell and B. LeBeau, "Student eligibility for a free lunch as an SES measure in educational research," *Educational Researcher*, vol. 39, pp. 120-131, 2010.
- [90] M. W. Ohland, M. K. Orr, V. Lundy-Wagner, C. P. Veenstra and R. A. Long, "Viewing access and persistence in engineering through a socioeconomic lens: Chapter 8," in *Engineering and Social Justice: In the University and Beyond*, West Lafayette, IN, Purdue University Press, 2012.

- [91] S. Aud, W. Hussar, M. Planty, T. Snyder, K. Blanco, M. Fox, L. Frohlich, J. Kemp and L. Drake, "The condition of education 2010 (NCES: 2010-028," National Center for Education Statistics, Institute of Education Sciences, US Department of Education, Washington, DC, 2010.
- [92] A. Owens, "Neighborhoods and schools as competing and reinforcing contexts for educational attainment," *Sociology of Education*, vol. 83, no. 4, pp. 287-311, 2010.
- [93] Advisory Committee on Student Financial Assistance, "The rising price of inequality: How inadequate grant aid limits college access and persistence," Department of Education, Washington, DC, 2010.
- [94] R. Rumberger and G. Palardy, "Does segregation still matter? The impact of student composition on academic achievement in high school," *Teachers College Record*, vol. 107, no. 9, pp. 1999-2045, 2005.
- [95] D. Hossler and F. Stage, "Family and high school experience influences on the post-secondary educational plans of ninth-grade students," *American Educational Research Journal*, vol. 29, no. 2, pp. 425-451, 1992.
- [96] W. Kimbrough and S. Harper, "African American men at historically Black colleges and universities: Different environments, similar challenges," in *African American Men in College*, San Francisco, CA, Jossey-Bass, 2006, pp. 189-209.
- [97] J. Ainsworth, "Why does it take a village? The mediation of neighborhood effects on educational achievement," *Social Forces*, vol. 81, no. 1, pp. 117-152, 2002.
- [98] L. Mendez, K. Raffaele, M. Howard and J. Ferron, "School demographic variables and out-of-school suspension rates: A quantitative and qualitative analysis of a large, ethnically diverse school district," *Psychology in the Schools*, vol. 39, no. 3, pp. 259-278, 2002.
- [99] J. Hoffman, "The impact of student cocurricular involvement on student success: Racial and religious differences," *Journal of College Student Development*, vol. 43, no. 5, pp. 712-739, 2002.
- [100] B. Munro, "Dropouts from higher education: Path analysis of a national sample," *American Educational Research Journal*, vol. 18, no. 2, pp. 133-141, 1981.
- [101] J. Zheng, K. Saunders, M. Shelley II and D. Whalen, "Predictors of academic success for freshmen residence hall students," *Journal of College Student Development*, vol. 43, no. 2, pp. 267-283, 2002.

- [102] F. Haemmerlie and R. Montgomery, "Gender differences in the academic performance and retention of undergraduate engineering majors," *College Student Journal*, vol. 46, no. 1, pp. 40-45, 2012.
- [103] J. Williams and M. Luo, "Understanding first-year persistence at a micropolitan university: Do geographic characteristics of students' home city matter?," *College Student Journal*, vol. 44, no. 2, 2010.
- [104] C. Mills, J. Heyworth, L. Rosenwax, S. Carr and M. Rosenberg, "Factors associated with the academic success of first year health science students," *Advances in Health Sciences Education*, vol. 14, no. 2, pp. 205-217, 2009.
- [105] A. Astin, W. Korn and K. Green, "Retaining and satisfying students," *Educational Record*, vol. 68, no. 1, pp. 36-42, 1987.
- [106] J. Fleming, "Who will succeed in college? When the SAT predicts Black students' performance," *Review of Higher Education*, vol. 25, no. 3, pp. 281-296, 2002.
- [107] R. Wolfe and S. Johnson, "Personality as a predictor of college performance," *Educational and Psychological Measurement*, vol. 55, no. 2, pp. 177-185, 1995.
- [108] T. Brown and R. Zwick, "Application of hierarchical linear modeling to a predictive validity study of college admissions tests," in *Paper presented at the annual meeting of the National Council on Measurement in Education*, San Francisco, CA, April 2006.
- [109] Louisiana Tech University, "Common Data Set 2014-2015," [Online]. Available: http://www.latech.edu/ir/assets/cds_2014-2015.pdf.
- [110] B. Pivik, "Comparing the ACT and SAT," Peterson's, A Nelnet Company, 25 August 2016. [Online]. Available: <https://www.petersons.com/college-search/comparing-act-sat.aspx>.
- [111] A. Braunstein, M. Lesser and D. Pescatrice, "The business of student retention in the post September 11 environment- Financial, institutional, and external influences," in *Proceedings of the Northeast Business & Economics Association*, 2006.
- [112] V. Lundy-Wagner, C. Veenstra, M. Orr, N. Ramirez, M. Ohland and R. Long, "Gaining access or losing ground? Socioeconomically disadvantaged students in undergraduate engineering, 1994-2003," *Journal of Higher Education*, vol. 85, no. 3, pp. 339-369, 2014.

- [113] D. Montgomery, E. A. Peck and G. Vining, *Introduction to linear regression analysis - Fifth Edition*, Hoboken, NJ: John Wiley and Sons, Inc., 2012.
- [114] D. Thistlewaite and D. Campbell, "Regression-discontinuity analysis: An alternative to the ex post facto experiment," *Journal of Educational Psychology*, vol. 51, pp. 309-317, 1960.
- [115] Institute of Education Sciences, "Postsecondary Education Evidence Review Protocol," 2013. [Online]. Available: <http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=242>.
- [116] J. Calcagno and B. Long, "The impact of postsecondary remediation using a regression discontinuity approach: Addressing endogenous sorting and noncompliance (An NCPR working paper)," National Center for Postsecondary Research, 2008.
- [117] A. Owen, "Grades, gender, and encouragement: A regression discontinuity analysis," *Journal of Economic Education*, vol. 41, no. 3, pp. 217-234, 2010.
- [118] S. Sun and W. Pan, "Investigating the accuracy of three estimation methods for regression discontinuity design," *The Journal of Experimental Education*, vol. 81, no. 1, pp. 1-21, 2013.
- [119] W. Trochim, *Research design for program evaluation: The regression discontinuity approach*, Newbury Park, CA: Sage, 1984.
- [120] H. Levin and J. Calcagno, "Remediation in the community college: An evaluator's perspective," *Community College Review*, vol. 35, no. 3, pp. 181-207, 2008.
- [121] W. Shadish, T. Cook and D. Campbell, *Experimental and quasi-experimental designs for generalized causal inference*, New York: Houghton-Mifflin, 2002.
- [122] D. Campbell and M. Russo, *Social experimentation*, Thousand Oaks, CA: Sage, 1999.
- [123] B. Moss and W. Yeaton, "Shaping policies related to developmental education: An evaluation using the regression-discontinuity design," *Educational Evaluation and Policy Analysis*, vol. 28, no. 3, pp. 215-229, 2006.
- [124] T. Melguizo, J. Bos, F. Ngo, N. Mills and G. Prather, "Using a regression discontinuity design to estimate the impact of placement decisions in developmental math," *Society for Research on Educational Effectiveness*, 2011.

- [125] J. Robinson, "Evaluating criteria for English learner reclassification: A casual approach using a binding-score regression discontinuity design with instrumental variables," *Educational Evaluation and Policy Analysis*, vol. 33, no. 3, pp. 267-292, 2011.
- [126] D. J. Hosmer and S. Lemeshow, *Applied logistic regression - Second Edition*, New York: Wiley, 2000.
- [127] C. J. Peng and T. So, "Logistic regression analysis and reporting: A primer," *Understanding Statistics*, vol. 1, no. 1, pp. 31-70, 2002.
- [128] J. Brown, G. Halpin and G. Halpin, "Relationship between high school mathematical achievement and quantitative GPA," *Higher Education Studies*, vol. 5, no. 6, 2015.
- [129] P. Lam, D. Doverspike and R. Mawasha, "Predicting success in a minority engineering program," *Journal of Engineering Education*, pp. 265-267, July 1999.
- [130] M. Orr, N. Ramirez and M. Ohland, "Socioeconomic trends in engineering: enrollment, persistence, and academic achievement," in *American Society for Engineering Education Annual Conference and Exposition*, Vancouver, Canada, June 2011.
- [131] D. Elster, "First-year students' priorities and choices in STEM studies- IRIS findings in Germany and Austria," *Science Education International*, vol. 25, no. 1, pp. 52-59, 2014.
- [132] M. Ensminger, C. Forrest, A. Riley, M. Kang, B. Green, B. Starfield and S. Ryan, "The validity of measures of socioeconomic status of adolescents," *Journal of Adolescent Research*, vol. 15, no. 3, pp. 392-419, 2000.
- [133] D. Gujarati, "Multicollinearity: What happens if the regressors are correlated?," in *Basic Econometrics -Fourth Edition*, McGraw-Hill.
- [134] G. M. Mullet, "Why regression coefficients have the wrong sign," *Journal of Qualitative Technology*, vol. 8, pp. 121-126, 1976.
- [135] S. Hahler and M. Orr, "Background and demographic factors that influence graduation: A comparison of six different majors," in *Frontiers in Education Conference*, El Paso, TX, 2015.
- [136] S. Tross, J. Harper, L. Osher and L. Kneidinger, "Not just the usual cast of characteristics: Using personality to predict college performance and retention," *Journal of College Student Development*, vol. 41, pp. 323-334, 2000.

- [137] ACT, Inc., "The Condition of College and Career Readiness 2016," 2016. [Online]. Available: http://www.act.org/content/dam/act/unsecured/documents/state19_Louisiana_Web_Secured.pdf.
- [138] K. Walker, "Research brief: Advantages and disadvantages of weighted grading," The Principals Partnership and Union Pacific Foundation.
- [139] B. Schneider, M. Carnoy, J. Kilpatrick, W. Schmidt and R. Shavelson, "Estimating causal effects sing experimental and observational designs," Governing Board of the American Educational Research Association Grants Program, Washington, DC, 2007.
- [140] D. Rubin, "Assignment to a treatment group on the basis of a covariate," *Journal of Educational Statistics*, vol. 2, no. 1, pp. 1-26, 1977.
- [141] P. Martorell and I. McFarlin, "Help or hindrance? The effects of college remediation on academic and labor market outcomes," *Society for Research on Educational Effectiveness*, 2009.
- [142] R. Jacob, P. Zhu, M. Somers and H. Bloom, "A practical guide to regression discontinuity," MDRC: Building Knowledge to Improve Social Policy, 2012.
- [143] Z. Zhang and R. RiCharde, "Prediction and analysis of freshman retention," AIR 1998 Annual Forum Paper, Minneapolis, MN, 1998.
- [144] J. Gayles, "Race, late bloomers and first-year GPA: Predicting beyond the freshman year," *Education Research Quarterly*, vol. 36, no. 1, pp. 13-29, 2012.
- [145] M. B. Paulsen and E. P. St. John, "Social class and college costs: Examining the financial nexus between college choice and persistence," *Journal of Higher Education*, vol. 73, no. 2, pp. 189-236, 2002.
- [146] J. King, *Crucial choices: How students' financial decisions affect their academic success*, Washington, DC: American Council on Education, 2002.
- [147] Louisiana Tech University, "Fact Book 2010-2014," Office of Institutional Research, 2016. [Online]. Available: http://www.latech.edu/ir/fact_book_2014/.
- [148] National Center for Education Statistics, "Condition of Education 2014," U.S. Department of Education, May 2014.
- [149] C. Belfield and P. Crosta, "Predicting success in college: The importance of placement tests and high school transcripts," *CCRC Working Paper No. 42*, Feb. 2012.

- [150] N. Peek, D. Arts, R. Bosman, P. von der Voort and N. de Keizer, "External validation of prognostic models for critically ill patients required substantial sample sizes," *Journal of Clinical Epidemiology*, vol. 60, no. 5, pp. 491-501, 2007.
- [151] F. Harrell, *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*, New York: Springer, 2001.